

**Executive Summary and Discussion of the  
Vagus Nerve Stimulation (VNS) Therapy Depression Indication  
Clinical Data (Updated to Include Information from Deficiency Letter Response)  
Prepared By: Richard L. Rudolph, MD, Vice President, Clinical and Medical Affairs and  
Chief Medical Officer, Cyberonics, Inc.**

## **1. Introduction**

### **1.1. Overview**

#### **1.1.1. Organization of Executive Summary**

This Executive Summary and Discussion of the Technical Section of the application consists of four parts: (1) an introduction, (2) a summary of the data supporting the effectiveness of VNS therapy for treatment-resistant depression (TRD), (3) a summary of the data supporting the safety of VNS therapy for TRD, (4) an evaluation of the risks and benefits of VNS therapy for TRD and overall conclusions.

#### **1.1.2. Background**

VNS therapy consists of intermittent stimulation (typically 30 seconds on, five minutes off) of the left cervical vagus nerve delivered via the VNS Therapy System. The principal components of the VNS Therapy System are an implantable pulse generator and lead, and an external programming system. In the United States, the VNS Therapy System is approved as an adjunctive therapy in reducing the frequency of seizures in adults and adolescents over 12 years of age with partial onset seizures which are refractory to antiepileptic medications. This application provides evidence to support the safety and effectiveness of VNS therapy for an additional indication as an adjunctive long-term treatment of chronic or recurrent depression for patients over the age of 18 who are experiencing a major depressive episode that has not had an adequate response to two or more adequate antidepressant treatments.

#### **1.1.3. VNS Therapy Depression Indication Clinical Development**

The development program for the depression indication conducted under Investigational Device Exemption (IDE) G980099 comprises two clinical studies, D-01 and D-02. Study D-01 is an open-label, uncontrolled feasibility study of VNS therapy in 60 subjects. Study D-02 is the pivotal study in this application. It consists of two phases, an acute, randomized controlled phase

and a long-term extension phase. In the acute phase, 235 subjects were implanted with the VNS Therapy System and then randomized to receive active VNS therapy (stimulation on) or sham stimulation (stimulation off) for 10 weeks. This 10-week period followed a 2-week post-operative recovery period during which stimulation remained off for all study subjects. In the long-term phase of the D-02 study, all subjects (ie, both the active phase acute treatment group and the sham-control group) who elected to continue past the acute phase received active VNS therapy. The Sponsor submitted a protocol for a third study, D-08 (intended as a post-marketing study), under the IDE, but withdrew it prior to study initiation.

This application also contains results from several clinical trials not conducted under IDE G980099: studies D-03, D-04, D-06, and selected add-on mechanism of action studies that examined various biological effects of VNS therapy in D-01 or D-02 study subjects. Study D-03 is an ongoing, open-label, uncontrolled post-marketing trial being conducted in Europe; D-03 safety data and preliminary efficacy data are included in this application. Study D-04 is an observational study of subjects not receiving VNS therapy who are well matched to the D-02 subjects for baseline demographic and disease characteristics. The D-04 subjects received standard-of-care treatment for depression. The outcomes for these subjects are used in this application as a control with which to compare the outcomes of the study subjects from the D-02 long-term phase. Study D-06 is an ongoing, open-label, uncontrolled feasibility trial to evaluate the safety and effectiveness of VNS therapy in subjects with rapid cycling bipolar disorder; it is being conducted under an investigator-initiated IDE, G000275. D-06 safety data are included in this application for completeness, but the D-06 subject population is distinct from the treatment-resistant depression subject population studied under IDE G980099. Consequently, pooling D-06 safety data (other than data related to the implant procedure itself) with those from the D-01, D-02, and D-03 studies would be inappropriate.

There are two additional protocols that were prepared for the pre-marketing clinical development of VNS therapy for the depression indication. No results from these protocols are included in this application for the following reasons. Study D-05 was not an actual clinical study; it was a videotape assessment of the D-02 study subjects used as part of the process to ensure adequate interrater reliability for the depression assessments. Study D-07 is an investigator-initiated IDE study that has not been initiated.

## 1.2. Unmet Medical Needs in the Management of TRD

Depression is a very common disorder that is most often chronic or recurrent in nature. In the United States, the lifetime prevalence of major depressive disorder (MDD), the type of depression studied under IDE G980099, is approximately 16%, and the 12-month point prevalence of MDD is approximately 7%.<sup>1</sup> Depression is associated with significant adverse consequences for the patient, patient's family, and society. Among the consequences of depression are functional impairment, impaired family and social relationships, increased mortality from suicide and comorbid medical disorders, and patient and societal financial burdens. Depression is the fourth leading cause of worldwide disability and is expected to become the second leading cause by 2020.<sup>2</sup>

There are currently three major treatment modalities for which there is substantial evidence of effectiveness in the treatment of a major depressive episode: pharmacotherapy with antidepressant drugs (ADDs), specific forms of psychotherapy, and electroconvulsive therapy (ECT). ADDs are the usual first line treatment for depression. Clinical trials have demonstrated efficacy for a number of pharmacologic classes of ADDs. Commonly the initial drug selected is a selective serotonin reuptake inhibitor (SSRI) such as fluoxetine (Prozac), or another of the newer ADDs such as venlafaxine (Effexor). Several forms of psychotherapy are used to treat depression. Among these, there is good evidence for the efficacy of cognitive behavior therapy and interpersonal therapy, but these treatments are used less often than are ADDs. Phototherapy is an additional treatment option that may be appropriate monotherapy for mild cases of depression that exhibit a marked seasonal pattern. Physicians usually reserve ECT for treatment-resistant cases or when they determine a rapid response to treatment is desirable.

Many patients do not respond to initial antidepressant treatment. Currently there are no treatments with an FDA-approved indication specifically for the treatment of such non-responders. Broadly speaking, physicians generally use one or more of the following strategies to treat patients who do not respond adequately to initial antidepressant treatment: (1) switching to an alternative first-line ADD, (2) switching to a second-line ADD (for example, a tricyclic ADD such as desipramine), (3) adding psychotherapy, a second ADD, or an augmentation agent. Augmentation agents are drugs that are not generally considered to have significant antidepressant activity when administered alone, but they can enhance the effectiveness of an ADD when they are administered in combination with the ADD. Augmentation agents include drugs such as lithium, triiodothyronine, or atypical antipsychotic drugs such as olanzapine. Additional options for treatment-resistant patients, especially for patients who fail on the above

alternatives, include monoamine oxidase inhibitors and ECT. For treatment-resistant cases that exhibit a marked seasonal pattern, adding phototherapy to pharmacotherapy may also be an option.

Despite the widespread availability of these treatment modalities, it is estimated that 10% to 20% of patients do not respond to treatment.<sup>3</sup> Even among patients who do respond, many do not respond completely, ie, they do not achieve symptom remission. Such partial responders remain at substantial risk for suicide,<sup>4</sup> future recurrences of full syndromic depression,<sup>5</sup> and significant functional impairment.<sup>6</sup> Moreover, there is little published evidence that any of the treatment strategies described above produces effective long-term control of depression in patients who fail to respond to initial antidepressant treatment. Furthermore, many treatments used for patients who do not respond at all or only respond partially to the first or second attempt at antidepressant therapy are poorly tolerated and/or are associated with significant toxicity. For example, tricyclic antidepressant drugs often cause anticholinergic effects and weight gain leading to premature discontinuation of therapy, and they can be lethal in overdose (a significant problem in depressed patients). Lithium is the augmentation strategy with the best published evidence of efficacy (although there are few published studies documenting long-term effectiveness), but lithium has a narrow therapeutic index that makes it difficult to administer; among the risks associated with lithium are renal and thyroid toxicity. Monoamine oxidase inhibitors are prone to produce an interaction with certain common foods that results in hypertensive crises. Even selective serotonin reuptake inhibitors can rarely produce fatal reactions in the form of a serotonin syndrome. ECT too is associated with significant risks: long-lasting cognitive impairment following ECT significantly limits the acceptability of ECT as a long-term treatment for depression. Therefore, there is a compelling unmet need for well-tolerated and effective long-term or maintenance treatments for patients who do not respond fully or who do not sustain a response to first-line antidepressant therapies.

### **1.3. Rationale For Developing VNS Therapy as a Treatment for Depression**

The initial rationale for investigating VNS therapy as a potential treatment for depression was derived from three observations. First, during the clinical trials that evaluated VNS therapy as a treatment for epilepsy there were reports that some patients experienced an improvement in mood. In some cases these improvements appeared to exceed what would be expected based on the amount of improvement observed in the patients' epilepsy. Second, physicians prescribe antiepileptic drugs as mood stabilizers in bipolar disorder and as augmentation agents to enhance

the effectiveness of antidepressant drugs. Third, ECT, which is often regarded as the most efficacious acute treatment for depression, exerts a marked anticonvulsant effect.

Subsequently several other findings have added to the rationale underlying VNS therapy's investigational use for the treatment of depression. Harden and Elger have both reported improvements in the depression scores of patients receiving VNS therapy for the treatment of epilepsy based on the use of standardized depression rating scales; similar to the anecdotal observations described above, these improvements were not dependent on improvement in the patients' seizures.<sup>7, 8</sup> Studies examining the central nervous system effects of vagus nerve stimulation also contribute to the rationale for VNS therapy's use as an antidepressant treatment. Based on the known projections of the vagus nerve within the brain and preliminary work examining the effect of vagus nerve stimulation on neurotransmitters, it appears that stimulating the vagus nerve produces effects on norepinephrine and serotonin, the two principal neurotransmitters implicated in the therapeutic effect of antidepressant drugs.<sup>9</sup> Additionally, neuroimaging studies show that VNS therapy modulates activity in regions of the brain believed to be involved in mood regulation, including limbic structures and specific cortical areas such as the orbitofrontal cortex. Finally, Krahel recently demonstrated that VNS therapy produces effects in the rat forced swim test (FST) paradigm that are comparable to those of the tricyclic antidepressant drug desipramine. The FST is a behavioral model commonly used to identify potential antidepressant drugs. (A "white paper" on neuroimaging findings relevant to depression and VNS and a report on the FST study are included in Section 9 of this PMA Supplement.)

## **2. Effectiveness**

### **2.1. Basis for Demonstrating the Effectiveness of VNS Therapy for the Treatment of Depression**

The demonstration of the effectiveness of VNS therapy for the treatment of depression is based primarily on analyses comparing the efficacy results from the D-02 study subjects during 12 months of VNS therapy plus standard-of-care treatment with the efficacy results from the D-04 study subjects during 12 months of only standard-of-care treatment (no VNS therapy). Additional supportive evidence for the antidepressant effectiveness of VNS therapy derives from the efficacy results of the D-02 acute study randomized controlled phase, the durability of the responses demonstrated during the long-term phase of both the D-02 and D-01 studies, and an indirect comparison of the D-02 study 12-month outcomes with the outcomes from a matched group of subjects in a multicenter study of electroconvulsive therapy (ECT) conducted by Columbia University.

## **2.1.1. Summary of the D-02 and D-04 Study Designs and the Statistical Plan for Comparing the D-02 and D-04 Efficacy Results**

### **2.1.1.1. D-02 Study Design**

The D-02 study is the pivotal study in this application. It consists of two phases, an acute, randomized controlled phase followed by a long-term extension phase. Subjects selected for the D-02 study were adults who met Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) criteria for a major depressive episode (MDE); subjects could have underlying major depressive disorder or bipolar disorder. The subjects' MDE had to be chronic or recurrent. In addition, subjects had to have failed in the current episode a minimum of two--but not more than six--adequate antidepressant treatments from different categories. The adequacy of prior treatment was determined using a modified version of a published scoring system developed to assess the adequacy of antidepressant treatments; a prior treatment was judged to be adequate if it received an antidepressant resistance rating of  $\geq 3$  in this system (referred to in this submission as the Antidepressant Treatment History Form guidelines). The subjects also had to have a baseline score of at least 20 on the 24-item Hamilton Rating Scale for Depression (HRSD-24).

Subjects who met these and the additional enrollment described in the protocol received a VNS Therapy System implant and then were randomized to have stimulation turned on or left off for the 12-week acute treatment phase. The acute phase consisted of a two-week recovery period following surgery during which stimulation remained off for all subjects, followed by a two-week stimulation adjustment period for those subjects randomized to active VNS therapy (the sham-control group underwent identical procedures as did the active VNS therapy group but had their output current set at 0 mA), followed by an eight-week period of treatment at the level of stimulation determined during the aforementioned stimulation adjustment period. Various clinical assessments, including depression ratings using standardized scales, were performed throughout the acute phase. Both the subjects and the study personnel performing the depression ratings were blinded as to treatment group assignment. The protocol-specified primary efficacy outcome for the acute phase was the percentage of treatment responders determined from the HRSD-24 ratings; response was defined *a priori* as a 50% or greater improvement in HRSD-24 score from baseline. The protocol permitted subjects to continue antidepressant treatments they were receiving prior to the study, but subjects could not increase antidepressant medication doses or add new antidepressant drugs.

Long-Term Extension Phase: All D-02 subjects who completed the acute phase, both those in the acute VNS therapy group and those in the sham-control group, were eligible to continue into the

long-term extension phase of the D-02 study. During this phase, all subjects received active VNS therapy. For the acute phase sham control subjects (also referred to as the delayed treatment group for the long-term phase), the first 10 weeks of the extension phase included stimulation parameter adjustments and weekly or every other week clinic visits and assessments identical to those experienced by the VNS therapy group during the acute phase. Otherwise, the protocol specified monthly clinic visits for both groups through 12 months of active VNS therapy. Various assessments, including depression ratings, were performed throughout this period. During the long-term extension phase, investigational site programmers were allowed to adjust stimulation parameters as clinically indicated. Additionally, concomitant antidepressant treatments could be added, removed, or adjusted as clinically indicated.

#### **2.1.1.2. D-04 Study Design**

The D-04 study is an observational study of subjects experiencing an MDE. The D-04 study was originally conceived to collect long-term clinical and health economic outcomes from patients receiving standard antidepressant therapies (ie, no VNS) for the treatment of depression that was similar in severity and level of treatment resistance to those of the D-02 subjects. The D-04 protocol had the same enrollment criteria as regards chronicity or recurrence of depression, prior treatment failures, and severity of depression as those described above for the D-02 study. Because the study was observational in nature, the protocol did not specify specific therapies for the treatment of depression; rather the physician managing the study subject's depression selected therapy according to his or her clinical judgment. The physician had the ability to select any antidepressant therapy available, except investigational treatments. Thus, antidepressant therapy in the D-04 study comprised "standard of care" treatment (also known as "treatment as usual"). The entire range of treatment options available for the D-04 subjects was also available to the D-02 subjects as concomitant treatment to their VNS therapy. Similar to the selection of treatments for the D-04 subjects, the selection of the concomitant treatments for the D-02 subjects was at the discretion of the physician managing the subject's depression. Thus subjects in both the long-term D-02 extension and the D-04 study received standard-of-care treatment; only the D-02 subjects also received VNS therapy.

In the D-04 study, the investigators performed various assessments, including depression ratings, at quarterly intervals. The specific battery of assessments differed somewhat from those obtained in the D-02 study, but evaluations based on the Inventory of Depressive Symptomatology Self Report (IDS-SR), HRSD-24, Clinical Global Impressions (CGI), and Medical Outcomes Survey 36-item Short Form Health Survey (SF-36) were common to both studies. The IDS-SR is a 30-

item, multi-dimensional patient self-report rating of the symptoms of mood and depression. This well-validated scale is becoming more widely used in clinical studies for assessing depression symptoms. The HRSD is the rating scale most widely used in clinical trials to assess symptoms of depression. The HRSD is a multi-dimensional observer-rated scale for assessing overall depression severity. There are several versions, including ones with 28, 24, 21, and 17 items. The 28-item version of the scale was administered in the D-02 and D-04 studies, but per protocol, only the first 24 items (HRSD-24) were used for rating purposes. The CGI includes two 7-point scales completed by the clinician rater to assess a subject's condition regarding severity of illness (CGI-S) and global improvement (CGI-I). The CGI was developed by the National Institute of Mental Health to provide a standardized assessment with clinically relevant anchors; it is one of the most widely used brief assessment tools in psychiatry. The SF-36 is a widely-used quality of life scale that assesses multiple domains.

#### **2.1.1.3. D-02/D-04 Comparison Statistical Analysis Plan**

A prospective statistical plan was formulated to compare the efficacy outcomes during 12 months of treatment for the D-02 study subjects (VNS Therapy plus standard-of-care) with those of the D-04 study subjects (standard-of-care only) for two reasons: (1) the results from the acute phase of the D-02 study were clinically meaningful but not statistically significant and (2) the acute D-02 sham-control group crossed over into VNS therapy at the end of the D-02 acute phase. This plan was submitted to the Food And Drug Administration September 3, 2002. The primary objective of the D-02/D-04 comparison statistical analysis plan was to demonstrate a difference in the improvement in depressive symptomology during 12 months of VNS therapy plus standard-of-care treatment (D-02 subjects) compared with 12 months of standard-of-care treatment without VNS therapy (D-04 subjects). The plan specified the primary population for efficacy analysis to be an "evaluable" population. The definition of this population is described in 2.2.1.

The plan further specified that the primary efficacy analysis was a repeated measures linear regression analysis of the raw IDS-SR scores during the first 12 months of active VNS therapy plus standard-of-care treatment (D-02 subjects) or during the first 12 months of standard-of-care treatment alone (D-04 subjects). The primary outcome in this analysis was the mean difference between the D-02 and D-04 groups in the estimated mean change from baseline in IDS-SR scores. This analysis uses all available scores. The repeated measures linear regression analysis incorporated a propensity score to control for differences in measured baseline covariates between the D-02 and D-04 study subjects. The IDS-SR scores were chosen for the primary efficacy analysis rather than the HRSD-24 scores because the linear regression analysis requires

multiple post-baseline observations and the only HRSD-24 observations available for the D-04 subjects were a baseline and a 12-month observation.

The plan also specified a large number of secondary efficacy analyses based on the IDS-SR, HRSD-24, CGI, and SF-36 assessments. Among these analyses were the determinations of the proportion of subjects achieving response and complete response (also referred to as remission). For the IDS-SR and HRSD-24, response was defined as a 50% or greater decrease from baseline score. For the CGI, response was defined as a rating of 1 (very much improved) or 2 (much improved) on the global improvement item. Complete response for the IDS-SR rating was defined as a score = 14. Complete response for the HRSD-24 rating was defined as a score = 9. An additional important feature of the analysis plan involved the calculation of the D-02 subjects' baseline depression scores. For the subjects originally assigned to the acute VNS therapy group, the baseline score was a mean of the scores obtained at the two pre-implant baseline assessments. For the D-02 subjects originally assigned to the sham-control group, the baseline score was a mean of the scores obtained at the last two assessments before the initiation of stimulation, ie, at the end of the acute treatment phase. This was a conservative approach that removed from the assessment of subject outcome any contribution from a pre-VNS therapy placebo response.

### **2.1.2. Valid Scientific Evidence for Determining Effectiveness**

This section describes how the D-02 and D-04 study protocols and the prospective statistical analysis plan for comparing the results from these two studies satisfy the requirement for the determination of the effectiveness of a medical device defined in 21 CFR Part 860.7 of the Code of Federal Regulations. 21 CFR Part 860.7 requires the Food and Drug Administration to rely only upon "valid scientific evidence to determine whether there is reasonable assurance that the device is safe and effective." 21 CFR Part 860.7 further states that the valid scientific evidence for the determination of effectiveness "shall consist principally of well-controlled investigations," which are defined as study plans and study reports that include five elements. Each of these five elements is described below along with a discussion of how the D-02 and D-04 protocols, the D-02/D-04 comparison statistical plan, or the actual D-02/D-04 comparative statistical analyses contained in this application satisfy each element.

#### **1) A clear statement of the study objective**

The statistical plan included a clear statement of the objective.

**2) A method for selecting study subjects that**

**a) Provides adequate assurance that the subjects are suitable for the purpose of the study and provides diagnostic criteria of the condition to be treated.**

The D-02 and D-04 protocols specified the enrollment criteria for the subjects. The principal enrollment criteria relevant to the proposed indication for use were as follows:

- ?? The subjects had to be diagnosed as being in a major depressive episode according to DSM-IV criteria ascertained from a structured interview (SCID).
- ?? The subjects' depression had to be chronic ( $\leq$  two years' duration) or recurrent (= four lifetime episodes).
- ?? The subjects had to exhibit a failure to respond to a minimum of at least two adequate (adequacy was defined in the protocol using specific criteria) antidepressant treatments from two different treatment categories.
- ?? The subjects had to have at least a minimum level of depression as defined by a score on the HRSD-24 of 20 or greater.

The D-02 and D-04 protocols were identical as regards the above criteria. Additionally, the D-02 protocol (but not the D-04 protocol) required subjects to have failed at least six weeks of psychotherapy. The impact, if any, of this additional criterion would have been to make the D-02 subjects more resistant to treatment than were the D-04 subjects. The protocols included clear diagnostic criteria for the fundamental disorder being studied, a major depressive episode. The criteria were derived from DSM-IV, which is the standard means for diagnosing a major depressive disorder and the same approach used in clinical antidepressant drug trials. Furthermore the protocols required that the presence of the DSM-IV criteria be ascertained via a structured interview, a more rigorous standard than is used in many antidepressant drug trials.

**b) Assigns subjects to test groups in such a way as to minimize any possible bias.**

**AND**

**c) Assures comparability between test groups and control groups of pertinent variables and use of concomitant therapy.**

Assuring comparability between test and control groups for pertinent demographic and disease characteristics is critical to the determination of whether differences in outcomes between the groups can be attributed to the test device. Although random treatment assignment is the optimum means to assure such comparability, 21 CFR 860.7 does not specifically require random treatment assignment. The D-02 and D-04 protocols used identical subject inclusion criteria for the key disease characteristics. The D-04 study was conducted at 13 investigational sites, 12 of which were also D-02 sites. The similarities in the key inclusion criteria and study sites provide a basis to expect that the demographic and disease characteristics of the D-02 and D-04 subjects

would be comparable, and thereby bias between test groups should be minimized. Furthermore, although investigators did not use random treatment assignment to determine specific study enrollment, there is no evidence that there was any selection bias in this process. There were no specific decision criteria for determining into which study to enroll a subject. Because the enrollment period for the two studies only partially overlapped, only 49 D-02 subjects could have enrolled in the D-04 study and only 10 D-04 subjects could have enrolled in the D-02 study. It is reasonable to assume the primary motivation for the 49 D-02 subjects to enroll in the D-02 study was for the opportunity to receive the investigational treatment. The primary reasons the 10 D-04 subjects opted for the D-04 study over the D-02 study were patient decision (7/10; subjects were concerned about surgery) or the presence of a D-02 protocol exclusion criterion related to the medical suitability of VNS therapy for the patient (3/10). The results of the analyses conducted to examine the comparability of the D-02 and D-04 subject groups confirm that the two groups are indeed comparable.

Assuring comparability between test and control groups for the use of concomitant therapies that could affect outcome is also critical to the determination of whether differences in outcomes between the groups can be attributed to VNS therapy. Both the D-02 and D-04 subjects received standard-of-care treatment. The protocols placed few limitations on the range of treatment options allowed. This was necessary because the subjects had failed many previous therapies. D-02 and D-04 subjects had the same range of concomitant treatment options available to them. Additions, removals, and adjustments in concomitant antidepressant therapies were permitted in both the long-term phase of the D-02 study and the D-04 study.

Consequently, an obvious critical issue for the interpretation of the results of the D-02/D-04 comparative analysis is the analysis of the comparability of the concomitant treatments. The results of this analysis show that concomitant antidepressant treatment additions or dosage increases occurred more frequently in D-02 non-responders and D-04 subjects than they did in the D-02 responders. Furthermore, when the primary efficacy analysis comparing the D-02 and D-04 outcomes was repeated with censoring of the D-02 data after the point at which significant changes in concomitant antidepressant therapy occurred, the VNS therapy group still demonstrated more improvement than did the D-04 standard-of-care group.

### **3) An explanation of the methods of observation and recording of results**

An explanation of the methods of observation and recording of results utilized is contained in the protocols for the D-02 and D-04 studies.

#### **4) A comparison of the results of treatment with a control that permits quantitative evaluation**

The determination of the effectiveness of VNS therapy for the treatment of depression is based primarily on a comparison of the results from quantitative rating scales used to compare outcomes during 12 months of VNS therapy plus standard-of-care treatment (study D-02) with the outcomes during 12 months of standard-of-care treatment alone (study D-04). Thus the determination of the effectiveness of VNS therapy is based on comparison to an active control. An active control is the most appropriate control to determine the effectiveness of VNS therapy for the proposed indication. It provides a more reliable determination of effectiveness than does the no treatment condition or historical control allowed in 21 CFR 820.7. Furthermore, although a placebo (ie, sham treatment) control was used for the acute phase trial, it appears that three months is too short a duration to permit separation sufficient to reach statistical significance between VNS therapy and the sham control. On the other hand, the use of a placebo control in association with no change in concomitant antidepressant treatment for more than 3 months in patients as ill as were the D-02 subjects is probably ethically unacceptable.

#### **5) A summary of the methods of analysis and an evaluation of the data**

The statistical plan for the D-02/D-04 comparison provides a prospective analysis plan. The D-02/D-04 study report included in this submission provides the evaluation of the data derived from the D-02 and D-04 studies.

### **2.2. D-02/D-04 Comparative Analysis**

#### **2.2.1. Efficacy Analysis Population**

The statistical plan specified that the primary population for the efficacy analyses would be an “evaluable population.” The definitions for the D-02 and D-04 evaluable populations are shown below, along with the number of subjects who qualified for each of the two evaluable populations.

**Table 1**  
**Criteria for Inclusion in Evaluable Population and Number of Subjects Qualifying**

<b>Study</b>	<b>Subjects Included</b>	<b>Number of Subjects</b>
<b>D-02</b>	<p>Subjects were included in the evaluable population if they:</p> <ul style="list-style-type: none"> <li>?? Met the intent-to-treat criteria, ie, were implanted, randomized, and met the acute phase continuation criteria</li> <li>?? Met protocol D-02 inclusion criteria 1, 2, 3 and 5<sup>?</sup></li> <li>?? Completed two baseline HRSD<sub>24</sub> assessments</li> <li>?? Received VNS therapy during the acute phase (if in the treatment group)</li> <li>?? Had no stimulation after day 35 through acute phase exit (if in the sham-control group [delayed treatment])</li> <li>?? Completed an HRSD<sub>24</sub> assessment at V8 or V9</li> <li>?? Had an acute exit HRSD<sub>24</sub> score of 18 or greater (if in the delayed treatment group)</li> <li>?? Received VNS therapy during the long-term phase</li> <li>?? And met at least one of the following criteria: <ul style="list-style-type: none"> <li>o Completed at least one HRSD<sub>24</sub> assessment post-acute phase exit</li> <li>o Were considered a treatment failure any time during the study</li> </ul> </li> </ul>	<b>(N=205)</b>
<b>D-04</b>	<p>Subjects were included in the evaluable population if they:</p> <ul style="list-style-type: none"> <li>?? Met Protocol D-04 inclusion criteria 1, 2, 3 and 5*</li> <li>?? Completed baseline IDS-SR assessment</li> <li>?? Completed at least one IDS-SR assessment post-baseline</li> </ul>	<b>(N = 124)</b>

For the D-02 group, 205 of the 233 implanted subjects who entered the D-02 long-term phase met criteria for the evaluable population. The reasons why subjects were excluded from the evaluable population are shown below.

---

<sup>?</sup> The protocol D-02 and D-04 inclusion criteria referred to here are as follows:

- 1 - Patient is diagnosed with a major depressive episode according to DMS-IV diagnosis criteria derived from the SCID.
- 2 - Patient is in a chronic (= two years) current MDE and/or has had a history of recurrent MDEs (at least four lifetime MDEs including the current MDE).
- 3 - Patient has not had an acceptable clinical response due to failure (resistance) with at least two treatments from different treatment categories during the current MDE.
- 5 - Patient has a score = 20 on the 24-item Hamilton Rating Scale of Depression.

**Table 2**  
**Subjects Who Were NOT Evaluable During the D-02 Long-Term Phase**

	<b>Delayed-Treatment</b>	<b>Treatment</b>	<b>Total</b>
<b>Reasons</b>	<b>Number of Subjects</b>	<b>Number of Subjects</b>	
No efficacy (assessments) data collected at any long-term visit	0	4	4
Did not meet acute phase continuation criteria	0	3	3
Did not have an acute exit HRSD <sub>24</sub> score = 18 (if in the delayed treatment [sham-control] group)	21	0	21
<b>Total</b>	<b>21</b>	<b>7</b>	<b>28</b>

The most common reason (21 of 28 cases) for excluding subjects from the D-02 evaluable population was that the subjects had an HRSD-24 score of less than 18 after exiting the acute phase of the D-02 study. These subjects were not suitable for the efficacy analyses because they demonstrated a placebo response prior to VNS therapy and therefore did not meet the protocol requirement for study continuation necessary for the evaluation of efficacy.

For the D-04 group, 124 of the 138 enrolled subjects met criteria for the evaluable population. Eleven (11) subjects signed a consent form but withdrew before completing the baseline assessments, and three subjects were excluded because they did not have post-baseline assessments.

### **2.2.2. Baseline Characteristics of Subjects**

Although the D-02 and D-04 groups were not created by random treatment assignment, it was expected that they would be very well-matched for baseline patient and disease characteristics because study subjects in both groups were required to meet enrollment criteria that were identical for the key variables and because the D-04 subjects were enrolled mainly at a subset of the D-02 investigational sites. The comparisons of the baseline characteristics for the evaluable populations confirm that the two treatment groups were indeed well matched.

More than 20 baseline characteristics were compared. Of these, only three showed statistically significant, or even marginally significant, differences between the groups: ethnic origin, prior ECT exposure, and the distribution of the number of lifetime episodes of depression. As both groups comprised at least 90% Caucasians, differences in ethnic origin are unlikely to be meaningful. D-02 subjects had more prior ECT exposure than did D-04 subjects, suggesting that

the D-02 group might be more treatment-resistant. Of the D-02 subjects, 53% received ECT at sometime during their lifetime and 35% received ECT in the current depressive episode prior to study enrollment. For D-04 subjects, the corresponding percentages were 26% and 12%, respectively ( $p < 0.001$  for both between-group comparisons). The D-02 and D-04 groups had a similar percentage of subjects with a history of five or fewer lifetime episodes of depression, but the D-02 group had a higher percentage of subjects with 6 to 10 lifetime episodes (D-02, 27%; D-04, 15%) and the D-04 group had a higher percentage of subjects with more than 10 lifetime episodes (D-02, 9%; D-04 26%;  $p < 0.001$  for the difference in overall distribution of prior episodes between the groups). This suggests the D-04 group might be more treatment-resistant. Thus prior ECT exposure and number of lifetime depressive episodes seem to offset each other in terms of any possible influence on the level of treatment resistance of the study groups. Prior use of specific classes of antidepressant drugs in the current episode was compared descriptively; statistical significance was not tested. The two groups were generally comparable in the percentage of subjects who had received various classes of antidepressant drugs in the current episode before study enrollment. Where moderate differences existed, the D-02 group always exhibited greater use of the class. This again suggests that the D-02 group might on the whole be more treatment-resistant than is the D-04 group. Finally, as regards the number of failed adequate prior antidepressant treatments in the current episode – the sole predictor of response to VNS therapy observed in the D-01 feasibility study – each group had a mean of 3.5 such treatments.

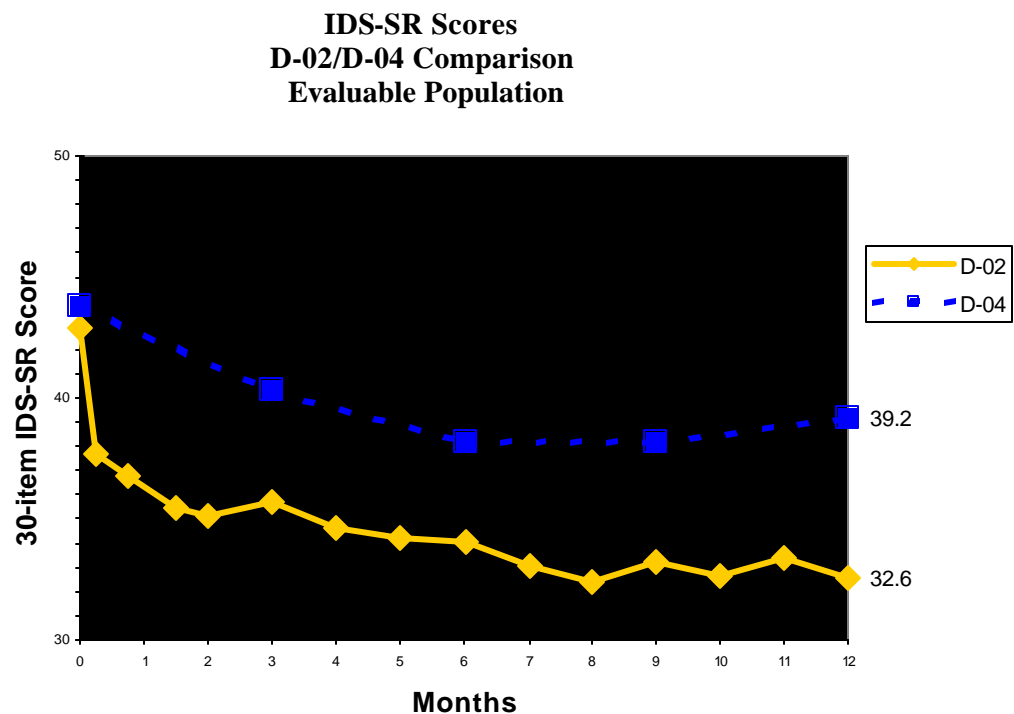
Thus baseline patient and disease characteristics were extremely well matched in the two subject groups. Additionally, a propensity adjustment strategy was incorporated into the primary efficacy analysis. Propensity scores were included as a covariate in the primary analysis to control for the effect of any differences between the groups in measured baseline characteristics. The propensity score did not have a significant effect on the analysis results ( $p = 0.831$ ). Thus, baseline patient and disease characteristics do not explain the differences in the outcomes between the D-02 and D-04 subjects.

### **2.2.3. Primary and Key Secondary Efficacy Analyses**

The primary efficacy analysis for comparing the results of VNS therapy plus standard-of-care treatment vs. standard-of-care treatment alone was a repeated measures linear regression analysis of IDS-SR raw scores over 12 months of treatment. The primary outcome in this analysis was the mean difference between the D-02 and D-04 groups in the estimated mean change from baseline in IDS-SR scores.

In this analysis, the VNS therapy group (D-02) showed greater improvement than did the standard-of-care treatment group (D-04) at a very robust level of statistical significance ( $p < 0.001$ ). The result is shown graphically in Figure 1. (As a point of clarity, note that for illustration purposes the graph shows the IDS-SR mean scores rather than the fitted regression lines. The  $p$ -value shown, however, is based on the statistical plan-specified primary efficacy analysis, ie, the repeated measures linear regression analysis.)

**Figure 1**



A large number of secondary efficacy analyses based on the IDS-SR, HRSD-24, and CGI-I ratings were also performed and provided confirmation of the result of the primary analysis. The VNS therapy group (D-02) showed more improvement than did the standard-of-care treatment group (D-04); almost all the differences were statistically significant, often robustly so. The results of the secondary analyses of the IDS-SR scores are shown in Table 3.

**Table 3**  
**IDS-SR Scores – D-02/D-04 Comparisons**  
**Evaluable Populations**

	<b>D-02</b>	<b>D-04</b>	<b>P-Value +</b>
<b>N</b>	180	112	
Baseline Average Raw Score (RS)	42.4	43.8	
12 Month Data			
Average RS	32.6	39.2	
Median RS	32	40	
Average Change	-9.8	-4.6	<0.001
LOCF Average Change <sup>++</sup>	-9.3 (N=204)	-5.0 (N=124)	<0.001
Median Change	-8.5	-3.5	
Avg. % Change	23.4	8.1	
Median % Change	20.6	7.9	
Response (% of Subjects)	22	12	0.029
LOCF Response (% of Subjects)	20 (N=204)	12 (N=124)	0.108
Complete Response (% of Subjects)	15	4	0.006
LOCF Complete Response (% of Subjects)	13 (N=204)	3 (N=124)	0.007

+Absence of a p-value indicates no statistical test was applied.

++LOCF stands for last observation carried forward and is defined in 2.2.4.

The results of the analyses of the HRSD-24 scores are shown in Table 4.

**Table 4**  
**HRSD<sub>24</sub> Scores – D-02/D-04 Comparisons**  
**Evaluable Populations**

		<b>D-02</b>	<b>D-04</b>	<b>P-Value</b>
<b>N</b>		180	104 <sup>1</sup>	
<b>Baseline Average</b>		27.9	27.8	
<b>12 Month Data</b>				
Average		19.6	22.8	
Median		19.5	23.5	
Average Change		-8.2	-4.9	0.006
LOCF Average Change <sup>++</sup>		-7.4 (N=205)	-4.9 (N=104)	0.04
Median Change		-7.5	-5.0	
Avg. % Change		29.6	16.6	
Median % Change		28.4	15.6	
Response (% of Subjects)		30	13	0.003
LOCF Response (% of Subjects)		27 (N=205)	13 (N=104)	0.011
Complete Response (% of Subjects)		17	7	0.031
LOCF Complete Response (% of Subjects)		17 (N=205)	7 (N=104)	0.059

<sup>1</sup> – 20 D-04 subjects did not have HRSDs performed at their 12-month visit; the 12-month HRSD was added after study initiation and several sites did not have IRB approval prior to subjects reaching the 12-month assessment in the study.

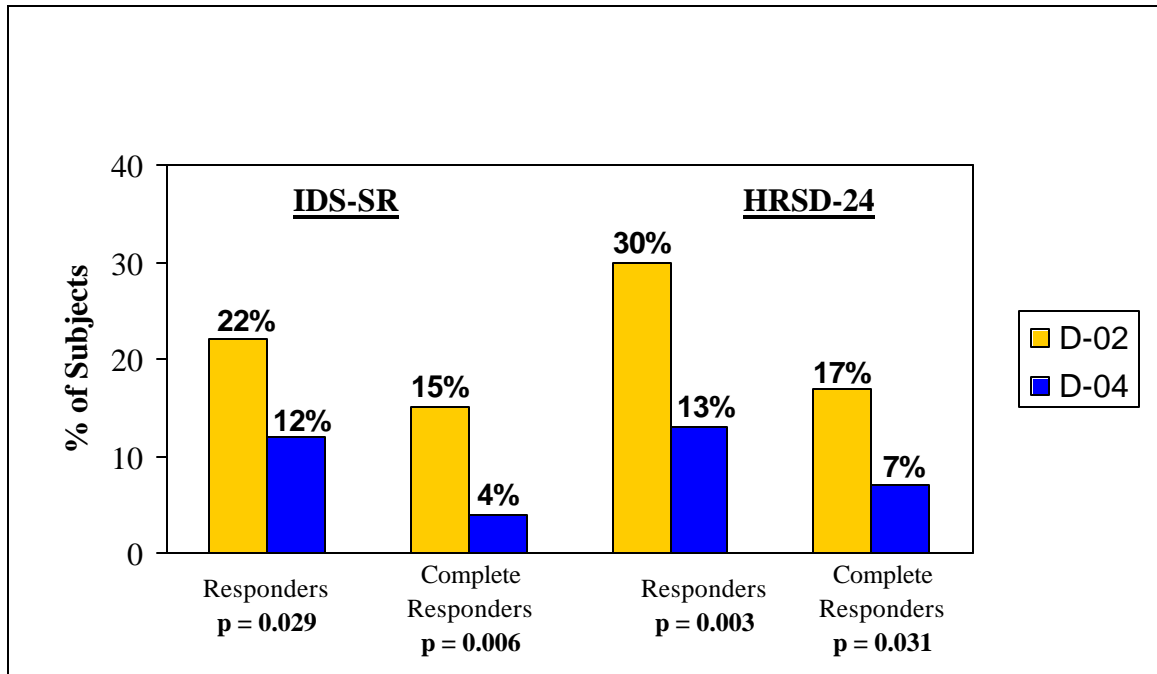
+Absence of a p-value indicates no statistical test was applied.

++LOCF stands for last observation carried forward and is defined in 2.2.4.

The decrease in HRSD-24 scores after 12 months was 8.2 for the VNS therapy group and 4.9 for the standard-of-care treatment group (p = 0.006). This result provides confirmation of the primary (IDS-SR linear regression) analysis result.

The percentages of subjects achieving a response or complete response based on the IDS-SR and HRSD-24 ratings are shown in Figure 2.

**Figure 2**  
**IDS-SR & HRSD<sub>24</sub> Response Rates and**  
**Complete Response Rates at 12 Months**  
**D-02/D-04 Comparisons**  
**Evaluable Population**  
**Observed Data**

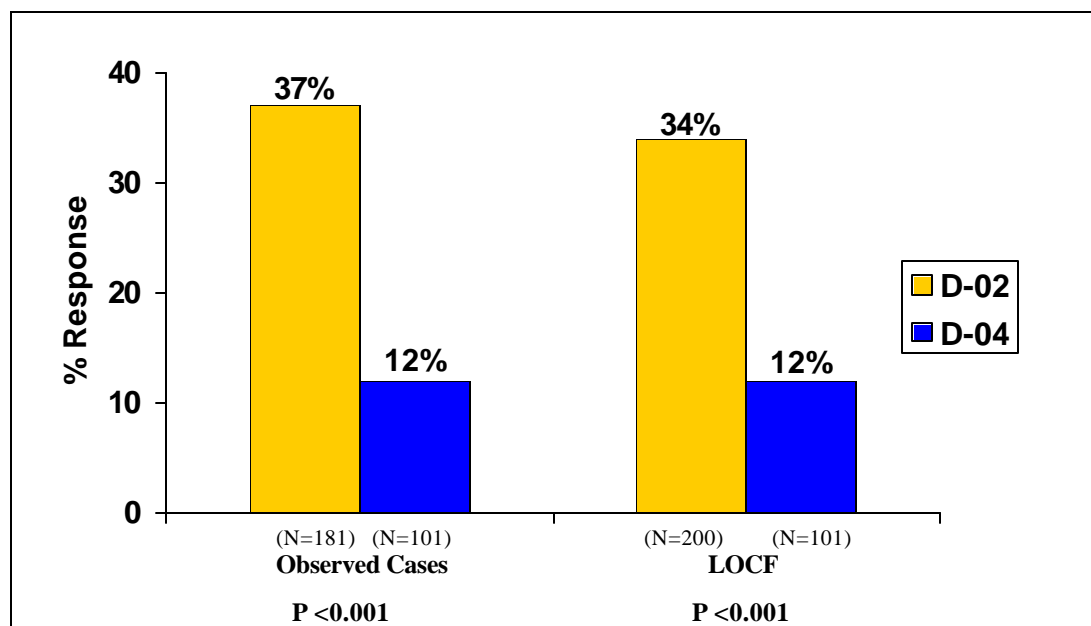


Note: N's available in Tables 3 and 4 above

The VNS therapy group had higher rates of response and complete response than did the standard-of-care treatment group; all the differences were statistically significant.

The percentages of subjects achieving a response based on the CGI global improvement assessment are shown in Figure 3.

**Figure 3**  
**Clinical Global Impressions – Improvement (CGI-I) Response Rates at**  
**12 Months, D-02/D-04 Comparisons**  
**Evaluable Population**



Three times as many subjects in the VNS therapy group as in the standard-of-care treatment group achieved a response based on CGI global improvement ratings (37% vs. 12% in the observed cases analysis;  $p < 0.001$ ).

*Sustained* response based on IDS-SR ratings was evaluated in an exploratory (ie, unplanned) analysis. Based on the definition recently proposed by Rush et al.,<sup>10</sup> sustained response was defined as a decrease in IDS-SR score of at least 50% from baseline at each of the last two quarterly evaluations (months 9 and 12). Sustained response was observed in 13% of the VNS therapy group and only 4% of the standard-of-care treatment group ( $p = 0.005$ ).

#### **2.2.4. Robustness of the Efficacy Results**

The statistical plan specified that the efficacy analyses would be performed on an evaluable population using observed data. To further evaluate the robustness of the results described in the preceding sections, several additional analyses were performed. These included statistical testing of the following: an intent-to-treat (ITT) analysis of the primary efficacy model, last observation carried forward (LOCF) analyses on the D-02 vs. D-04 secondary analyses, an analysis of the primary efficacy model limited to data obtained from the subset of D-02 sites also involved in the D-04 study. The purpose of these additional analyses was to demonstrate that the superior D-02

outcomes (1) were not altered if all D-02 implanted subjects were included in the primary analysis, (2) were not altered when missing observations were accounted for, and (3) were not explained solely by depression improvement in subjects from the D-02 sites that did not participate in the D-04 study.

An ITT analysis typically includes all randomized subjects regardless of whether they receive adequate treatment or continue for an adequate duration. Therefore this is a conservative approach to analyzing treatment outcomes that eliminates any potential bias that may arise by excluding certain subjects from the efficacy analysis. In an ITT analysis of the primary efficacy model (N = all 235 implanted D-02 subjects and all 127 D-04 subjects), the mean difference between the D-02 and D-04 groups in the estimated mean change from baseline in IDS-SR scores continued to show a superior outcome in the D-02 group at the same level of significance as shown in the evaluable population analysis ( $p < 0.001$ ). As an even more conservative modification of the ITT approach, a second ITT analysis was performed that excluded the four D-02 subjects who did not meet the acute continuation criteria (ie, they had a response prior to the initiation of stimulation) and the three D-04 subjects who did not have post-baseline assessments. Therefore this ITT analysis included 231 D-02 subjects and 124 D-04 subjects. The results from this second ITT analysis were also statistically significant at  $p < 0.001$ .

An LOCF analysis is a technique often used by statisticians to address the impact of missing data. It uses the last available observation for subsequent time points where data are missing. For treatments associated with an increasing effect over time, LOCF understates the true treatment effect. LOCF analyses were performed on all D-02/D-04 secondary comparisons, and statistical significance was maintained for all comparisons except for the IDS-SR evaluable response rates and HRSD-24 evaluable complete response rates; in these latter two analyses, the results approached statistical significance.

Where sites are not identical between two studies, the differential outcomes among these sites can contribute to biasing the outcome. Since the D-02 and D-04 studies had some different sites, this was a potential concern. Therefore results were examined from sites that were only involved in both the D-02 and D-04 studies. When the primary efficacy analysis (repeated measures linear regression analysis of IDS-SR raw scores) was performed using data only from those sites that participated in both studies, the D-02 group still demonstrated greater improvement in depression scores than did the D-04 group at a very robust level of statistical significance ( $p = 0.002$  in the evaluable population and  $p = 0.003$  in the ITT population). In addition, most of the secondary analyses of the IDS-SR and HRSD-24 data also demonstrated statistically significant advantages

for the D-02 group, despite the reduced sample size in the D-02 group. Therefore, the superior efficacy outcomes observed in the overall VNS therapy group (compared to the standard-of-care treatment group) cannot be attributed to a differential improvement between the subjects enrolled at D-02 sites that did not participate in the D-04 study and subjects enrolled at D-02 sites that did participate in the D-04 study.

In summary, several alternative analyses as described above were applied to the comparisons of the D-02 and D-04 efficacy outcomes. The objective of these alternative analyses was to determine if applying conservative statistical approaches to address the potential biases that might be inherent in the primary and secondary analysis strategies would alter the statistical significance or the conclusions of the D-02/D-04 comparisons. When the D-02 and D-04 results were compared with the conservative alternative analyses, the D-02 outcomes almost always remained statistically significantly superior to the D-04 outcomes. Thus the robust statistical significance demonstrated in the primary and secondary analyses was supported by more conservative supplemental analyses.

#### **2.2.5. Efficacy Results in Subgroup of Subjects with Bipolar Disorder**

The D-02 and D-04 protocols permitted the enrollment of subjects with either unipolar or bipolar (manic-depressive) depression. Of the evaluable populations, 10% (20/205) of the D-02 subjects and 12% (15/124) of the D-04 subjects had a diagnosis of bipolar disorder. Analyses were performed to compare the D-02 and D-04 efficacy outcomes in the bipolar disorder subgroups. Generally, the numerical results in the bipolar subjects were similar to those in the overall groups. For example, the 12-month HRSD-24 response rate in the D-02 evaluable population was 29% (5/17) for the bipolar subjects (vs. 30% for the overall D-02 evaluable population). Owing to the small sample sizes, however, differences between the D-02 and D-04 subjects were not statistically significant.

#### **2.2.6. Discussion of D-02/D04 Comparative Analysis Results**

The analyses comparing the outcomes of subjects with treatment-resistant depression treated for 12 months with VNS therapy and standard-of-care treatment with the outcomes of similar subjects receiving only standard-of-care treatment for 12 months show that VNS therapy is an effective antidepressant treatment. The results of the primary analysis, a repeated measures linear regression analysis of IDS-SR raw scores, showed that improvement was greater in the group receiving VNS therapy, and this result was highly statistically significant ( $p < 0.001$ ). The finding that the group receiving VNS therapy in addition to standard-of-care treatment had more

improvement than did the group receiving standard-of-care treatment alone was supported by the results from a large number of secondary analyses based on the IDS-SR ratings and ratings from a variety of other standard depression rating scales. For example, the proportion of subjects rated as responders to therapy for the group receiving VNS therapy plus standard-of-care treatment vs. the group receiving standard-of-care treatment alone was 22% vs. 12% ( $p = 0.029$ ) for the IDS-SR, 30% vs. 13% ( $p = 0.003$ ) for the HRSD-24, and 37% vs. 12% ( $p < 0.001$ ) for the CGI-global improvement item. Moreover, a variety of analyses showed that the demonstration of statistical significance was independent of the population chosen for analysis (evaluable population or intent-to-treat population) or the method used to select observations for analysis (observed data or last observation carried forward).

The chief limitation of the D-02/D-04 comparative analysis is that the data are derived from a comparison of well-matched subject groups treated with two different interventions rather than from a randomized, subject data set. The use of the D-04 subjects as a control for the D-02 subjects does, however, provide a high degree of confidence that the goal of a randomized controlled study is achieved. This goal is to minimize the various sources of potential bias that could provide an alternate explanation to a treatment effect of the investigational therapy as the reason that there is a difference in outcomes between two interventions.

The primary sources of bias that a randomized design is intended to mitigate can be conceptualized as differences between patients in a treatment and control group as regards:

?? patient characteristics

?? disease characteristics

?? concomitant treatment

What follows is a discussion of how 1) the results obtained from analyses comparing the D-02 and D-04 baseline demographic and disease characteristics, 2) the use of a propensity adjustment in the primary efficacy analysis, and 3) results from analyses comparing concomitant antidepressant treatments in the D-02 and D-04 subjects show that material differences between the D-02 and D-04 subjects in patient characteristics, disease characteristics or concomitant antidepressant treatments did not exist, and therefore the superior outcomes in the D-02 subjects can be attributed to VNS therapy. Additionally there is a discussion as to why it is unlikely that the improvement in depression scores observed over 12 months in the D-02 subjects is attributable to a placebo response.

### **2.2.6.1. Patient and Disease Characteristics**

The D-04 subjects, while not a randomized control group, were enrolled at a subset of the D-02 sites using selection criteria identical to those of the D-02 protocol as regards the critical inclusion criteria. Thus it was expected that the D-02 and D-04 subjects would be well matched in their baseline patient and disease characteristics. The analysis of a large number of patient and disease characteristics confirmed this to be the case: only three characteristics were shown to be statistically significantly different between the D-02 and D-04 evaluable populations used for the efficacy analyses. One of these, ethnic origin, is extremely unlikely to constitute a clinically meaningful difference as Caucasians constituted at least 90% of both the D-02 and D-04 groups. The other two statistically significant differences were the percentages of subjects who had ECT (higher in the D-02 group), either lifetime or in the current depressive episode prior to study enrollment, and the distribution of the number of lifetime episodes of depression (the D-04 group had a higher percentage of subjects with more than 10 lifetime episodes). These differences probably offset each other as regards their contribution, if any, to the level of treatment resistance of the subjects. Importantly, the sole predictor of response to VNS therapy in the D-01 feasibility study, which was the number of failed adequate antidepressant treatments in the current depressive episode, was represented equally in the D-02 and D-04 groups: each group had a mean of 3.5 adequate failed antidepressant treatments in the current episode prior to enrollment. Thus the D-02 and D-04 subjects appear to be essentially identical as regards observed prognostically important covariates.

As an additional measure to ensure that differences in baseline patient and disease characteristics do not explain differences in the outcomes between the VNS therapy group and the standard-of-care group, a propensity adjustment strategy was incorporated into the primary efficacy analysis comparing the D-02 and D-04 subjects' outcome. In this two-stage strategy, first a propensity score was calculated. The propensity score represents an individual's probability of being in the D-02 study vs. the D-04 study conditional on the individual's covariate values. In the second stage, the propensity score was incorporated in the primary efficacy analysis to control for the effect of any differences between groups in measured baseline patient and disease characteristics. For the D-02 comparison with D-04, the propensity score did not contribute to the statistical significance of the primary outcome ( $p = 0.831$ ). This is a further indication that measured baseline differences in patient and disease characteristics do not explain the difference in the outcome between the D-02 and D-04 subjects in the primary efficacy analysis.

Propensity adjustment cannot, however, account for differences in unobserved (or unknowable) prognostic factors. Nevertheless, it is a reasonable assumption that the groups do not differ significantly on unobserved prognostic factors (if any exist). This assumption is reasonable because the measured baseline covariates included most (if not all) known prognostic factors, the differences between the D-02 and D-04 subjects in these covariates were negligible (further confirmed by the propensity adjustment result), and the D-02 and D-04 groups were both quite large (205 and 124 subjects respectively). The patient and disease covariates that were not recorded at baseline, such as family predisposition for affective disorder, have not been demonstrated to be significant predictors of treatment response. Furthermore, the large treatment cell sizes provide a high degree of assurance that any unobserved covariates would likely be equally distributed between the D-02 and D-04 groups, or would be so rare as to be of no consequence in their impact on the statistical significance of the results. It is difficult to imagine how the typical randomized controlled antidepressant drug trial with approximately 100 subjects (often fewer) per group would produce more well-matched treatment groups than are found in the D-02 and D-04 studies.

#### **2.2.6.2. Concomitant Antidepressant Treatments**

Both the protocol for the D-02 long-term phase and the protocol for the D-04 study permitted the use of antidepressant therapies (drugs and ECT) other than VNS therapy. Only D-02 subjects received VNS therapy. Therefore attribution of improvement in subjects' depressive symptomatology to VNS therapy requires both that D-02 subjects improve more than do D-04 subjects and that the use of the concomitant (ie, non-VNS) antidepressant treatments be no greater in the D-02 subjects than it is in the D-04 subjects. To the contrary, if the VNS therapy group demonstrates more improvement than does the standard-of-care treatment group, one would expect the use of concomitant antidepressant therapies to be less in the VNS therapy group. This is precisely what occurred.

ECT use was similar among the D-02 and D-04 subjects (7% and 6% respectively). On the other hand, antidepressant drug use was higher in D-02 subjects who were non-responders and D-04 subjects overall than it was in the D-02 subjects who achieved a response. During the 12 months of observation, 77% of the D-02 non-responders and 81% of all D-04 subjects either added a new antidepressant treatment or increased an existing antidepressant dose by an antidepressant resistance rating (ARR) level of one or more. By contrast, only 56% of the D-02 subjects who were responders to VNS therapy either added a new antidepressant treatment or increased an existing antidepressant dose by an ARR level of one or more ( $p < 0.01$  for comparison with D-02

non-responders and  $p < 0.001$  for comparison with D-04 subjects). Additionally, twice as many D-02 responders (44%) had no ARR changes or removed or decreased medications by at least one ARR level or were not taking medications as compared to non-responders (23%). These findings strongly support the conclusion that the VNS therapy responders derived their benefit from VNS therapy (or an augmentation effect of VNS therapy), rather than from the concomitant antidepressant treatments.

Even more persuasive in excluding an effect of the concomitant antidepressant treatments as the explanation for the superior outcomes in the VNS therapy group is the result that was obtained from repeating the primary repeated measures linear regression analysis of IDS-SR scores after censoring the D-02 subjects' scores for concomitant antidepressant treatment changes. In this alternative analysis, D-02 subjects' scores were censored when the addition of an antidepressant treatment or an increase in an existing treatment by an ARR level of one or more occurred. The analysis incorporated additions or changes in either antidepressant drugs or ECT. In other words, if a D-02 subject added or increased a concomitant antidepressant treatment, his or her subsequent IDS-SR scores were not used in this confirmatory repeated measures linear regression analysis. Instead the subject's last IDS-SR score before the concomitant antidepressant treatment change was used for subsequent assessment points (ie, a last-observation-carried-forward approach). Consequently this analysis removes from the VNS therapy group outcome any potential benefit from an addition or increase in concomitant antidepressant treatment. The approach is asymmetric as no censoring is performed on the D-04 data. It is an overly conservative approach because it penalizes VNS therapy by removing all incremental benefit that is attributable to VNS therapy after the censoring time point. Therefore, it understates the treatment effect of VNS therapy. For the asymmetric censored analysis, the result for the primary outcome (ie, the mean difference between the D-02 and D-04 groups in the estimated mean change from baseline in IDS-SR scores) is  $-0.183$  with a standard error of  $0.09$  (95% confidence interval of  $-0.37$  to  $0.00$ ), and the  $p$ -value is  $0.052$ . The estimated fourth quarter (endpoint) mean IDS-SR scores for the D-02 and D-04 groups were  $36.135$  and  $38.327$ , respectively, for a difference of  $-2.191$  over the 12 months.

Despite using this very conservative approach, the repeated measures linear regression analysis of IDS-SR scores was still marginally statistically significant ( $p = 0.052$ ). This constitutes very strong evidence that the superior outcome in the VNS therapy group is not due to concomitant antidepressant treatment. The 95% confidence interval ( $-0.37$  to  $0.00$  for the D-02/D-04 difference in IDS-SR score per month where negative values indicate D-02 superiority) shows

that the addition of VNS therapy alone (ie, without any additions or increases in concomitant antidepressant treatments) for less than one year of treatment was virtually always more effective than were *multiple* medication/ECT additions or increases over a full year of treatment using any currently available therapeutic option (standard-of-care).

#### **2.2.6.3. Why the Improvement Observed in the D-02 Subjects is Not Likely to Be Attributable to a Placebo Response**

Antidepressant treatment trials conducted in the general (ie, non-treatment-resistant) depressed population are often plagued by the presence of high placebo response rates (typically 30% or more during short-term trials). Several observations and findings, however, make it unlikely that a placebo effect accounts for the improvement observed in the D-02 subjects after 12 months of VNS therapy. Most importantly, there are consistent and robust statistically significant differences between the outcomes in the D-02 subjects and the outcomes in the D-04 subjects who are an active treatment control group. Additionally placebo response rates in the treatment-resistant depression (TRD) population are much lower than they are in the general depressed population. Thase and Rush have reported that placebo response rates in TRD are in the range of 0% to 10%.<sup>11</sup> The results observed in the D-02 sham-control group and the D-04 subjects provide strong confirmation for this statement. Furthermore, Quitken et. al. have shown that when placebo response occurs in depression, it is an early phenomenon that does not persist.<sup>12</sup> (Therefore, most of the improvement attributable to a placebo effect in the subgroup of the D-02 cohort who were in the original D-02 sham-control group and later received VNS therapy was presumably removed when the subjects' baseline scores were reset at the end of the acute phase of the D-02 study.)

To further evaluate the potential contribution of a placebo effect to the outcomes in the D-02 subjects, the Sponsor performed a comparative analysis of the 12-month outcomes in the subgroups of evaluable subjects from the D-02 and D-04 studies who were in a chronic (ie, at least 2 years in duration) major depressive episode at the time of study enrollment. Owing to the low overall placebo response rate in TRD cited by Thase and Rush and the transient nature of placebo response in depression reported by Quitkin et. al., it is particularly unlikely that these chronic depression subgroups would exhibit a placebo response after 12 months of treatment. Response and complete response rates, calculated from the HRSD-24 scores, were 29% (35/121) and 14% (17/121), respectively, for the D-02 subjects and 10% (7/72) and 3% (2/72), respectively, for the D-04 subjects. (Formal statistical testing was not performed on these results because the comparison was an unplanned exploratory analysis.) These results are similar to the results in the full D-02 and D-04 evaluable populations and provide additional evidence that the

improvements in depression scores observed in the D-02 subjects are attributable to VNS therapy and not to a placebo effect.

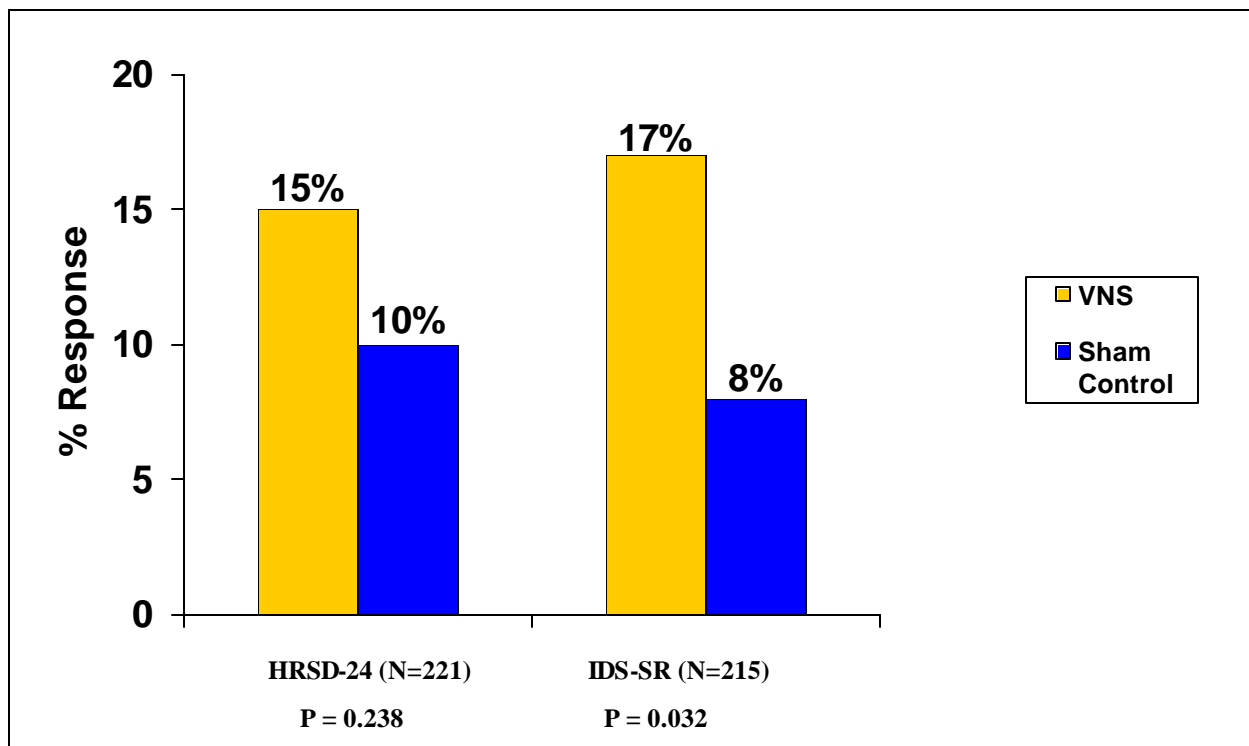
### **2.3. Supportive Effectiveness Data**

Several sources of evidence other than the D-02/D-04 comparison provide additional results supporting the effectiveness of VNS therapy as an antidepressant treatment.

#### **2.3.1 D-02 Acute Randomized Controlled Phase Study Results**

The primary efficacy outcome for the acute phase was the percentage of treatment responders determined from the HRSD-24 ratings; response was defined as a 50% or greater improvement in the HRSD-24 score from baseline to the end of the acute phase. This analysis was conducted on the statistical plan-defined evaluable population and based on observed (actual) data. A numerically higher percentage of subjects in the VNS therapy group than in the sham-control group (15% vs. 10%) were responders (Figure 4); the difference was not statistically significant ( $p = 0.238$ ). There was a consistent trend throughout the secondary efficacy outcomes for the VNS therapy group to show numerically greater improvement than did the sham-control group. These differences reached statistical significance in the analysis of the percentage of treatment responders determined from the IDS-SR ratings (where response was defined as a 50% or greater improvement in the IDS-SR score from baseline to the end of the acute phase) and in an exploratory analysis that used IDS-SR scores in a repeated measures linear regression analysis. In the IDS-SR responder analysis, 17% of the VNS therapy group were responders vs. 8% of the sham-control group ( $p = 0.032$ ; Figure 4). The results of the D-02 acute phase efficacy analyses summarized above provide supportive evidence for the effectiveness of VNS therapy as a treatment for depression.

**Figure 4**  
**D-02 Study Acute Phase**  
**Percent Responders at Visit 9**  
**Evaluable Population**  
**Observed Data**



### 2.3.2. Additional D-02 Long-Term Phase Study Results

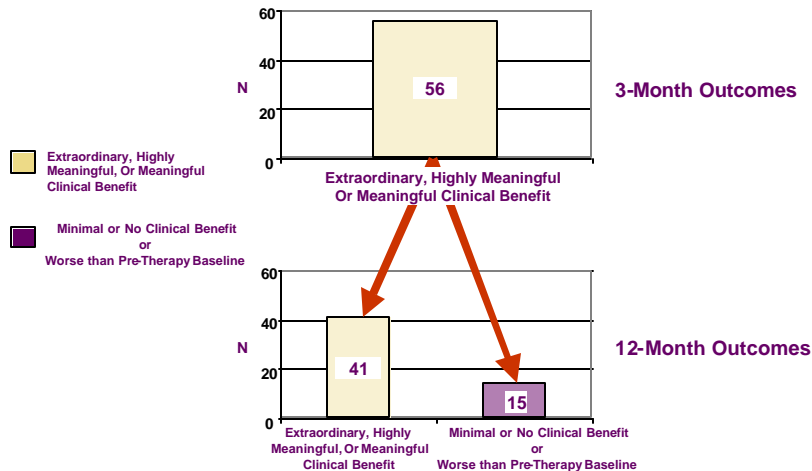
Results from additional analyses of the D-02 long-term phase HRSD-24 ratings also provide supportive evidence for the effectiveness of VNS therapy. These analyses used the repeated HRSD-24 observations during 12 months of VNS therapy to evaluate the durability of response to VNS therapy, ie, the degree to which improvement associated with VNS therapy was sustained. These analyses could not be performed on the data from the D-04 study because the D-04 study protocol only specified HRSD-24 ratings be collected at baseline and after 12 months of treatment during the first year of observation. Nonetheless as described in 2.2.6.3, even in the absence of a control, the presence of a sustained improvement during VNS therapy constitutes strong supportive evidence of VNS therapy's effectiveness because placebo response--and even antidepressant drug response--would be expected to diminish significantly and fairly rapidly over time in a patient population as treatment-resistant as are the D-02 subjects.<sup>12-16</sup>

Two types of analyses using HRSD-24 scores were performed to evaluate the durability of the improvements observed in the D-02 subjects who completed 12 months of VNS therapy. The

first analysis was a determination of the proportion of D-02 subjects who achieved a *sustained* response during 12 months of treatment. This outcome was identified in the D-02 long-term statistical plan as the key clinical endpoint for determining whether any improvements observed in the D-02 subjects were *clinically* significant. Sustained response was defined in the statistical plan based on the HRSD-24 ratings between study months 9 and 12, a period that included four HRSD-24 ratings. Sustained response was defined as an improvement from baseline in HRSD-24 score of at least 50% on at least one of these ratings plus an improvement of at least 40% on at least two other ratings. A group of nationally-recognized depression experts established *a priori* a success criterion of “approximately 30%” as a benchmark for a clinically significant sustained response rate. The sustained response rate for the D-02 subjects achieved this benchmark: the proportion of D-02 subjects who achieved a sustained response was 27%.

The second analysis used to determine the durability of the improvements observed in the D-02 subjects was a shift table analysis examining categories of improvement. Subjects who had paired observations after 3 and 12 months of VNS therapy (N=174) were categorized into three pre-specified categories based on their percentage change from baseline in HRSD-24 scores: 1) extraordinary (= 75% reduction) or highly meaningful (50% to 74% reduction) clinical benefit, 2) meaningful clinical benefit (25% to 49% reduction), 3) minimal or no clinical benefit (0% to 24% reduction) or worse than baseline. The durability of improvements in HRSD-24 scores after 3 months of VNS therapy was examined for each individual subject by comparing the subject’s 3-month status with his or her 12-month status. For example, a subject who exhibited highly meaningful clinical benefit (50% to 74% reduction in HRSD-24 score) after 3 months of VNS therapy and exhibited the same or greater reduction in HRSD-24 score after 12 months of VNS therapy can be considered to maintain their improvement, ie, demonstrate improvement that is durable. The results of this analysis show that most of the D-02 subjects maintained or increased their improvement: 34 of 56 subjects (61%) who exhibited extraordinary, highly meaningful or meaningful clinical benefit after 3 months of VNS therapy exhibited the *same* or *more* improvement after 12 months of VNS therapy. Additionally 41 of these 56 subjects (73%) maintained *at least* meaningful clinical benefit (Figure 5).

**Figure 5**  
**12-Month Outcomes of D-02 Subjects with Extraordinary, Highly Meaningful**  
**or Meaningful Clinical Benefit At 3-Months (N=56)**



? 73% (41/56) of subjects maintained at least a meaningful benefit.

The proportion of VNS therapy-treated subjects maintaining their response compares favorably to published rates associated with maintenance therapy following an acute response to ECT in patients who are arguably less treatment-resistant. For example, a seminal publication by Sackeim et al. reported that 24 weeks after response to ECT, 61% of patients receiving a combination of the tricyclic antidepressant nortriptyline plus lithium and only 40% of patients receiving nortriptyline alone had not relapsed; among patients treated with placebo, only 16% maintained their response for 24 weeks.<sup>16</sup> In a second published study, Sackeim et al. reported that relapse following ECT response was almost twice as likely among patients who were medication-resistant: only 32% of such patients maintained their response during the year following ECT.<sup>15</sup>

The clinically significant sustained response rate and the high rate of maintained responses during 12 months of VNS therapy observed in the D-02 subjects provide indirect – but strongly supportive – evidence of the effectiveness of VNS therapy.

### 2.3.3. D-01 Feasibility Study

#### 2.3.3.1. D-01 Study Design

The initial human trial to investigate the use of VNS therapy for the treatment of depression was the D-01 feasibility study. The D-01 study was an open-label, uncontrolled trial of VNS therapy

in adults who met DSM-IV criteria for a major depressive episode (MDE). The study subjects could have a diagnosis of major depressive disorder or bipolar disorder. In addition to the presence of a MDE, the protocol required that (1) the subjects' MDE was chronic or recurrent in nature, (2) the subjects had to have failed at least two adequate antidepressant treatments (based on the use of the Antidepressant Treatment History Form guidelines) from different categories of treatment in the current depressive episode, and (3) the subjects had a score of at least 20 on the 28-item Hamilton Rating Scale for Depression (HRSD-28) at the time of enrollment. Subjects who met enrollment criteria received a VNS Therapy System implant. The study had an acute treatment phase and a long-term follow-up phase.

The acute phase consisted of a two-week recovery period from the implant surgery during which stimulation remained off, followed by a two-week period during which investigator adjusted stimulation to the maximum tolerated level, followed by an eight-week period of treatment at the maximum tolerated level of stimulation. Various assessments, including depression ratings using standardized scales, were performed throughout this acute phase for comparison to baseline (pretreatment) scores. The protocol-specified primary efficacy outcome was the percentage of treatment responders, with response defined as a 50% or greater improvement from baseline in HRSD-28 score. Although the protocol permitted subjects to continue antidepressant treatments they were receiving prior to the study, subjects could not increase antidepressant medication doses or add new antidepressant drugs.

In the long-term follow-up phase, subjects who completed the acute phase and elected to continue treatment continued to receive VNS therapy and undergo periodic assessment. During this follow-up phase, the protocol permitted adjustments to both the VNS therapy stimulation parameters and to the concomitant antidepressant medications.

#### **2.3.3.2. D-01 Study Results**

Sixty (60) subjects were implanted. Sixty-five percent (65%) of the subjects were female. The mean age of the 60 subjects was 46.8 years. The mean length of their current depressive episode was 9.9 years, and lifetime 18.0 years. The mean number of treatments received by these subjects in their current depressive episode prior to study enrollment was 15.7; two thirds of the subjects had a prior history of receiving ECT sometime during their lifetime. As required by the protocol, all subjects failed at least two antidepressant treatments documented as adequate in dose and duration in the current depressive episode; 37 subjects (62%) failed four or more such treatments. The baseline disease characteristics of the D-01 study subjects indicate that they are an unusually treatment-resistant group of subjects.

Fifty-nine (59) of the 60 subjects were included in the analysis of efficacy; one subject whose HRSD-28 ratings improved during the two-week post-implant recovery period (when stimulation was still off) was excluded from the analysis. At the end of the acute treatment phase, 31% (18/59) of the evaluable subjects were responders based on the primary efficacy outcome, the HRSD-28 response rate. Fifteen percent (15%) or nine of the 59 evaluable subjects demonstrated a complete response, defined as a HRSD-28 score = 10. These response and remission rates are quite remarkable when one considers the average level of chronicity and prior treatment-resistance of the D-01 study subjects. The results from the secondary depression rating measures confirmed those from the HRSD-28.

The HRSD-28 response and complete response rates for subjects with assessments one year after implantation were 45% (25/55) and 27% (15/55) respectively. The response and complete response rates for subjects with assessments two years after implantation were 43% (18/42) and 21% (9/42) respectively. Even if one uses a conservative intent-to-treat approach to the analysis of the two-year results and assumes all subjects without a two-year assessment were treatment failures, the two-year response and complete response rates are 31% (18/59) and 15% (9/59) respectively, suggesting that the level of response observed after the acute treatment phase was well-maintained over a two-year period.

To evaluate the maintenance of response and complete response further for individual subjects, a shift table analysis was done similar to the one described above for the D-02 long-term phase. Subjects who had paired observations after 3 and 12 months of VNS therapy (N=55) were categorized according to the magnitude of the change in their HRSD-28 scores at the end of the acute treatment phase vs. the magnitude of the change in their HRSD-28 scores after 12 months of treatment (there were too few subjects with paired observations at 3 and 24 months to create meaningful shift tables for that pairing). The shift tables show that most of the D-01 subjects maintained or increased their improvement: 20 of 30 subjects (67%) who exhibited extraordinary, highly meaningful or meaningful clinical benefit after 3 months of VNS therapy exhibited the *same* or *more* improvement after 12 months of VNS therapy. Additionally 23 of these 30 subjects (77%) maintained *at least* meaningful clinical benefit. Although the uncontrolled study design used for the D-01 study precludes definitive conclusions about the efficacy of VNS therapy, the results summarized above – particularly the response and complete response rates observed after 12 and 24 months of treatment and the results of the shift table analyses – provide additional supportive evidence that VNS therapy is effective for the treatment of depression.

#### **2.3.4. A Comparison of the VNS Therapy Outcomes from the D-02 Study with Outcomes from a Multicenter Study of Electroconvulsive Therapy (ECT)**

Many psychiatrists consider ECT to be the most effective acute therapy available for the treatment of depression. Owing to poor patient acceptance and safety concerns (particularly adverse cognitive effects), however, physicians generally reserve the use of ECT for patients who do not respond to or cannot tolerate pharmacotherapy, situations in which a rapid response is deemed desirable, and certain other limited situations (eg, psychotic depression). These same concerns severely limit the use of ECT as a maintenance therapy. Consequently while some responders to an acute course of ECT receive longer-term maintenance ECT, most receive maintenance pharmacotherapy.

Prudic et al. recently completed an observational study of the effectiveness of ECT at seven community hospitals in the New York City metropolitan region.<sup>17</sup> One of the investigators (Harold Sackeim, Ph.D.) of that study defined a subset of the study subjects in an attempt to provide a cohort that was well-matched to the D-02 study subjects. This provided a unique opportunity to compare the long-term effectiveness of VNS therapy plus standard-of-care treatment (ie, the outcomes from study D-02) with the effectiveness of standard-of-care post-ECT maintenance treatment (the outcomes in the matched ECT cohort from the Prudic study).

Dr. Sackeim created the matched ECT cohort by excluding from the full data set subjects with the following characteristics that would have excluded the subjects from enrollment in the D-02 study:

1. Age less than 18 or greater than 72;
2. Baseline HRSD-24 score = 19;
3. Presence of psychotic symptoms;
4. Diagnosis of schizoaffective disorder.

The resulting matched ECT cohort comprised 172 of the 347 subjects in the original study population. The matched ECT cohort was similar to the D-02 long-term phase evaluable population (N=205) in sex distribution (65% female vs. 64% female, respectively), baseline mean HRSD-24 score (31 vs. 28), and mean age (48 years vs. 46 years). The matched ECT cohort had a greater proportion of subjects with a diagnosis of bipolar disorder than did the D-02 group (20% vs. 10%, respectively) and an older mean age of onset of mood disorder (29 years old vs. 22 years old, respectively). Overall, however, the D-02 group appears to comprise a more treatment-resistant population than does the matched ECT cohort because the D-02 group presented with a longer mean current episode duration than did the matched ECT cohort (50 vs. 11 months,

respectively), a greater proportion of subjects with a history of prior ECT (53% vs. 32%), more prior hospitalizations for depression (mean of 2.7 vs. 1.5), a longer mean duration of illness (26 years vs. 19 years), and more failed adequate treatments in the current episode (a mean of 3.5 vs. 1.7).

Dr. Sackeim determined response and remission rates for the matched ECT cohort. He defined response as a 50% or greater improvement from baseline in HRSD-24 score. He defined remission as a 60% or greater improvement from baseline to a score of 10 or less on the HRSD-24 rating. (The definition of complete response [remission] used in analyzing the D-02 long-term results was slightly different in that it used a cutoff score of 9 or less but no percentage change criterion.) In the matched ECT cohort, 58% of the subjects achieved response and 44% of the subjects achieved remission following acute treatment with ECT. Maintenance of remission, however, was poor among these subjects. Six months following the completion of the acute course of ECT, 41% of the matched ECT cohort met the criterion for response; only 20% met the criteria for remission. Thus, response and remission were not maintained after a successful acute course of treatment with ECT despite the availability and use of standard antidepressant treatments.

By contrast, response and remission rates, as determined from the HRSD-24 ratings, increased from the end of month 3 to the end of month 12 during continuous treatment with VNS therapy plus standard-of-care treatment. In the D-02 evaluable population, 14% (29/205) met the response criterion and 7% (14/205) met the remission criterion after 3 months of treatment. After 12 months of VNS therapy plus standard-of-care treatment, 27% (55/205) met the response criterion and 15% (30/205) met the remission criterion. (Note that the D-02 response and remission rates described above differ from those described earlier in 2.2.3, **Primary and Key Secondary Efficacy Analyses**, because the rates described here are derived from a more conservative analysis that more closely corresponds to the analysis done by Dr. Sackeim. In the earlier response and remission calculations, the rates were calculated based only on subjects with actual observations. Here rates are calculated by dividing the number of subjects meeting the response or remission criterion by the entire sample of 205.)

Substantial differences in the study designs, analysis methods, and subject populations of these two studies preclude any definitive conclusions. Nonetheless, the striking contrast between the *declining* response and remission rates during post-ECT standard-of-care maintenance treatment and the *improving* response and remission rates during maintenance VNS therapy plus standard-

of-care treatment provides additional supportive evidence of the long-term effectiveness of VNS therapy plus standard-of-care treatment.

#### **2.4. Clinical Significance of the Efficacy Results**

The efficacy outcomes described above suggest that VNS therapy will provide clinically significant benefits to patients with TRD. To understand the clinical significance of the efficacy results, they must be considered in the context of the study population. The subjects enrolled in the D-01 and D-02 trials represent patients whose disorder is analogous to terminal disorders such as end-stage congestive heart failure or metastatic cancer refractory to conventional therapy. Since only a relatively small percentage of such patients can generally be expected to respond to therapy (with the exception perhaps of cardiac transplant for end-stage congestive heart failure), one cannot readily determine clinical significance by examining mean outcomes in the total treated population. Rather, clinical significance is best determined by examining categorical outcomes, such as response rates. After 12 months of adjunctive VNS therapy in the D-02 study, response rates ranged from 22% to 37%, complete response (which corresponds to a patient being essentially completely well) rates ranged from 15% to 17%, and sustained response rates ranged from 13% to 27%. For the population studied, who had been ill for a mean of 50 months and who had received a mean of 10 medication treatments (including a mean of four failed adequate treatments) in their current episode of depression, these are clinically meaningful rates. The response rates above, defined by the standard definition of an improvement of at least 50% in depression score, may actually understate the clinical benefit for patients as chronically ill and treatment-resistant as were the D-01 and D-02 subjects. Such patients may experience clinically significant benefit from an improvement in depression score in the range of 25% to 49%. Based on this broader criterion, 100 of the 180 (56%) evaluable population D-02 subjects who completed the 12-month HRSD-24 assessment derived clinical benefit.

Yet another way to appreciate the clinical significance of the D-02 efficacy outcomes is in relation to the amount of difference typically observed between treatments in antidepressant drug trials. Based on meta-analyses of published antidepressant drug trials in adult outpatients with major depressive disorder, the Agency for Health Care Policy and Research's Clinical Practice Guideline for depression reported differences in response rates for drug-placebo trials of 20% for SSRIs and 21% for tricyclic antidepressants. Drug-drug differences in active treatment control trials of SSRIs and tricyclic antidepressants were negligible.<sup>18</sup> The differences in response rates between the two active treatment conditions described in this application (ie, adjunctive VNS

therapy in the D-02 study and standard-of-care treatment in the D-04 study) were similar to the drug-placebo differences--and far exceeded the drug-drug differences--cited above.

## **2.5. Effectiveness Conclusions**

?? The results from the comparison of depression ratings between D-02 study subjects (receiving VNS therapy plus standard-of-care treatment) and D-04 study subjects (receiving only standard-of-care treatment; no VNS therapy) over 12 months of observation provide valid scientific evidence of the effectiveness of VNS therapy for the adjunctive long-term treatment of chronic or recurrent depression for patients over the age of 18 who are experiencing a major depressive episode that has not had an adequate response to two or more adequate antidepressant treatments. Furthermore, the magnitude of the improvements observed in the D-02 subjects represents a clinically significant outcome.

- The primary efficacy analysis (a repeated measures linear regression analysis of IDS-SR scores) showed a highly statistically significant difference ( $p < 0.001$ ) in favor of the VNS therapy group.
- The secondary efficacy analyses also demonstrated statistically significant results in favor of the VNS therapy group vs. the standard-of-care only treatment group. These included the change from baseline to 12 months in the HRSD-24 score (difference of 3.3 points;  $p = 0.006$ ), IDS-SR response rates (22% vs. 12%;  $p = 0.029$ ), IDS-SR complete response rates (15% vs. 4%;  $p = 0.006$ ), IRS-SR sustained response rate (13% vs. 4%;  $p = 0.005$ ), HRSD-24 response rates (30% vs. 13%;  $p = 0.003$ ), and CGI improvement response rates (37% vs. 12%;  $p < 0.001$ ).
- Alternate statistical analysis techniques using intent-to-treat-analysis and last-observation-carried forward approaches demonstrated the robustness of the results described above. Additionally, a subset analysis limiting the D-02 data to only those investigational sites also participating in the D-04 study provided further evidence of the robustness of the primary analysis result.
- Extensive analyses of baseline patient and disease characteristics and concomitant antidepressant treatment use during the 12-month studies demonstrate that (1) differences between the D-02 and D-04 subjects in baseline patient and disease characteristics were few and do not account for the difference in efficacy outcomes between the subject groups and (2) differences in the use of concomitant

antidepressant treatments do not account for the difference in efficacy outcomes between the D-02 and D-04 subject groups.

?? Observations from several sources other than the D-02/D-04 study comparison provide additional supportive evidence of the effectiveness of VNS therapy as an antidepressant treatment.

- The comparison of outcomes in the VNS therapy group vs. the sham-control group in the D-02 acute randomized controlled phase showed a consistent trend for modestly greater improvement in the VNS therapy group with a statistically significant difference on a secondary outcome, the IDS-SR response analysis (17% vs. 8%;  $p = 0.032$ ).
- 27% of subjects completing 12 months of VNS therapy in the D-02 study were sustained responders (HRSD-24). 73% of the subjects who had paired observations after 3 and 12 months of VNS therapy and who achieved a level of meaningful clinical benefit or better after 3 months of VNS therapy maintained at least a level of meaningful benefit after 12 months of VNS therapy. In the D-01 feasibility study, 77% of the subjects who had paired observations after 3 and 12 months of VNS therapy and who achieved a level of meaningful clinical benefit or better after 3 months of VNS therapy maintained at least a level of meaningful benefit after 12 months of VNS therapy. Proportions of treatment-resistant subjects as high as these achieving sustained response and maintaining acute response are strongly suggestive of significant antidepressant efficacy.
- As demonstrated by the cohort of subjects from the Prudic community ECT study who were chosen to match the D-02 subjects for key characteristics, response and remission rates decline during post-ECT standard-of-care maintenance treatment. By contrast, in the D-02 study, response and remission rates increased from the end of month 3 to the end of month 12 during maintenance VNS therapy plus standard-of-care treatment.

?? The efficacy outcomes summarized above suggest that adjunctive VNS therapy will provide *clinically* significant benefits to patients with treatment-resistant depression. The response, complete response, and sustained response rates cited above indicate that a clinically significant percentage of study subjects--who on average had an illness of very long duration that had not adequately responded to a large number of standard treatment interventions--

experienced clinically meaningful improvement while receiving adjunctive VNS therapy. The differences in response rates between the adjunctive VNS therapy group (D-02) and the standard-of-care treatment group (D-04) are similar to drug-placebo differences--and far exceed drug-drug differences--reported in the medical literature from adult outpatient antidepressant drug trials.

### **3. Safety and Tolerability**

The initial safety and tolerability of the VNS Therapy System were established in the epilepsy clinical trials submitted in the epilepsy indication PMA application that was approved in 1997. The safety and tolerability of VNS therapy have subsequently been confirmed during commercial use of VNS therapy in more than 22,000 patients with epilepsy who comprise over 56,000 patient-years of experience. The VNS Therapy Systems and the range of stimulation parameters used in the depression clinical trials are the same as have been used in commercial experience. Consequently the principal objective for the safety evaluation in the depression clinical trials was to determine if the safety and tolerability of the VNS Therapy System differed from that observed in the epilepsy trials and commercial experience, i.e., to determine if the safety and tolerability of the VNS Therapy System were altered due to differences between the epilepsy and depression patient populations or their disease states.

#### **3.1. Safety and Tolerability of the VNS Therapy System in Patients with Epilepsy**

Experience from the epilepsy IDE clinical trials and over 56,000 patient-years of commercial use has established the VNS Therapy System as safe and well tolerated.

##### **3.1.1. Safety of Implant Procedure in Patients with Epilepsy**

The most commonly reported serious adverse events associated with implantation of the VNS Therapy System are infection and nerve injury. Infection necessitating explant of the VNS Therapy System was reported in 1% (5/454) of the patients in the IDE clinical trials and has been reported in 1% of commercial implants. The infection rate associated with VNS Therapy System implants appears to be similar or lower than those of similar implantable devices. Nerve injury usually manifests as left vocal cord paralysis or left facial paralysis. The rate of all nerve injuries reported in the IDE clinical trials was 1%; the rate reported in commercial use is less than 0.5%. These rates appear to be lower than the rate of nerve injury reported in the medical literature for carotid endarterectomy. During commercial use, 47 incidences of significant bradycardia or transient asystole during the initial intraoperative lead test have been reported (47/24,640: 0.19%). In some of these cases, the surgeon discontinued the implant procedure.

### **3.1.2. Overall Tolerability and Commonly Reported Adverse Events Associated with VNS Therapy in Patients with Epilepsy**

VNS therapy is very well tolerated by patients with epilepsy as demonstrated by extraordinarily high rates of treatment continuation in the IDE clinical trials: more than 95% of the study subjects continued to receive VNS therapy past 12 months. The most commonly reported adverse events in the high (therapeutic) stimulation groups in the two pivotal IDE clinical trials (studies E-03 and E-05) were voice alteration/hoarseness (39-73%), cough (12-53%), throat pain (7-42%), pain (34% in E-05), dyspnea (11-27%), and paresthesia (16-24%). Often these events are present only when the VNS Therapy System is stimulating the vagus nerve, and typically a patient reports these events much less frequently over time.

### **3.1.3. Less Commonly Reported But Potentially Serious Adverse Events Associated with VNS Therapy in Patients with Epilepsy**

Less commonly reported but potentially serious adverse events associated with VNS Therapy in patients with epilepsy are listed in 4.1 below, **Risks Associated with the Use of VNS Therapy in the Treatment of Depression.**

## **3.2. Safety Profile of VNS Therapy in Patients with Treatment-Resistant Chronic or Recurrent Depression**

The safety profile of VNS therapy observed in the studies conducted under IDE G980099, the D-01 and D-02 studies, is consistent with the known safety profile of VNS therapy that has emerged from the epilepsy IDE clinical trials and over 56,000 patient-years of commercial use. Based on the data submitted in this application, which include complete safety data obtained through the specified data cutoff dates for all D-01 and D-02 subjects implanted with the VNS Therapy System and serious adverse event data obtained through the specified data cutoff date for non-IDE D-03 subjects implanted with the VNS Therapy System, the use of VNS therapy in patients with treatment-resistant, chronic or recurrent depression appears to be associated with similar adverse effects and safety risks as does its use in patients with treatment refractory epilepsy.

### **3.2.1. Common Implant-Related Adverse Events**

Since all eligible D-01 and D-02 study subjects received VNS Therapy System implants, there was no control available for assessing whether or not an adverse event reported post-implant was related to the surgery. Therefore the determination of which adverse events reported by subjects were related to implantation was based on investigator assessment. The events reported as at least possibly related to implantation that occurred in at least 5% of the subjects who received VNS Therapy System implants in the D-01 or D-02 studies were device site pain, device site reaction, headache, incision pain, neck pain, pain, dyspepsia, dysphagia, nausea, abnormal

healing, hypertonia, hypesthesia, paresthesia, cough increased, dyspnea, laryngismus, pharyngitis, voice alteration, and incision site reaction. These events are expected consequences of surgery and are consistent with the labeling contained in the current VNS Therapy Pulse Model 102/102R Generator Physician's Manual. The implant-related events in study D-02 that occurred in 10% or more of subjects (device site pain, device site reaction, incision pain, dysphagia, hypesthesia, pharyngitis, voice alteration and incision site reaction) were analyzed to determine how long they persisted. Most of the individual incidences of these events resolved within 30 days. Hypesthesia (generally described as a localized numbness) and voice alteration, however, tended to be more persistent in a substantial number of individuals. For example, in 17 of 24 reports of implantation-related hypesthesia in the D-02 study the event continued beyond 3 months. Similarly, in 21 of 81 reports of voice alteration the event continued beyond 3 months. Hypesthesia would be an expected symptom of nerve injury during surgery. Its persistence in some subjects suggests the presence of a nerve injury that may be slow to heal or permanent. The persistence of voice alteration in some subjects is difficult to assess because it could represent surgical injury to the innervation of the larynx, but vagus nerve stimulation itself can cause voice alteration.

### **3.2.2. Common Stimulation-Related Adverse Events**

For the acute phase of the D-01 and D-02 studies, the respective protocols required all adverse events representing a new event or change from baseline, regardless of relationship to VNS therapy, to be reported. For the long-term phases, only events continuing from the acute phase, new events/prior events with increased severity that in the investigator's judgment were related to VNS therapy, and serious events regardless of relationship had to be reported. Consequently many, if not most, of the adverse events reported during the D-01 and D-02 studies are related to the subjects' underlying illness, comorbid illnesses, medications, or other intercurrent events and not to VNS therapy. The most informative analysis to determine which adverse events are related to stimulation is the comparison of the adverse events reported for the VNS therapy and sham-control groups during the D-02 study acute phase. Following a convention used in antidepressant drug trials to determine which adverse events are likely to be treatment-related, the Sponsor determined which of the adverse events reported during the D-02 acute phase (regardless of the investigators' judgment as to relationship to VNS therapy) occurred in at least 5% of the VNS therapy group and at a frequency at least twice that of the sham-control group. The events meeting these criteria are shown in Table 5 below.

**Table 5**  
**D-02 Acute Phase Treatment Group Adverse Events Occurring at  $\geq 5\%$  and  $\geq 2 \times$  Sham-Control Group**

<b>Adverse Event</b>	<b>Treatment Group N (%)</b>	<b>Sham-Control Group N (%)</b>
Neck Pain	25 (21%)	11 (10%)
Wound Infection	9 (8%)	2 (2%)
Dyspepsia	12 (10%)	6 (5%)
Vomiting	13 (11%)	6 (5%)
Cough Increased	35 (29%)	10 (9%)
Laryngismus	13 (11%)	2 (2%)

Two additional adverse events nearly met the criteria: dysphagia (21% of the VNS therapy group and 11% of the sham-control group), and voice alteration (68% of the VNS therapy group and 38% of the sham-control group). The above events, with the exception of wound infection, are all well-known effects of stimulating the vagus nerve and are all described in the current VNS Therapy Pulse Model 102/102R Generator Physician's Manual. Wound infection, an obvious risk of the implant procedure, is not likely to be related to stimulation, and its greater occurrence in the VNS therapy group probably represents an anomaly.

Another approach used to determine which of the commonly reported adverse events might be vagus nerve stimulation-related was to identify those events reported by the investigators during the acute phase in the D-02 study that the investigators judged were at least possibly stimulation-related and which occurred at a frequency of 10% or greater in the VNS therapy group. The following adverse events met the above criteria: neck pain (16%), dysphagia (13%), paresthesia (10%), cough increased (24%), dyspnea (19%), laryngismus (11%), voice alteration (55%). Note that the percentages cited differ from those in Table 5 because that table includes all reports of the event and the percentages immediately above are only for incidences judged by investigators to be stimulation-related. This second approach to identifying commonly reported adverse events associated with stimulation identifies two additional events not captured by the  $\geq 5\%$  /  $\geq 2 \times$  criteria: paresthesia and dyspnea. Again, both of these events are well-known effects of vagus nerve stimulation and are described in the current physician's manual.

The seven events identified above as stimulation-related and occurring at a frequency  $\geq 10\%$  in the VNS therapy group were analyzed further to determine how long they persisted. For most of the seven events, subjects were no longer reporting the event after 60 days. Laryngismus and

voice alteration, however, tended to be more persistent with 7 of 13 and 53 of 68 subjects, respectively, continuing to report the event more than 3 months after the onset of VNS therapy.

### **3.2.3. Study Discontinuations Due to Adverse Events**

Overall, the continuation rates for VNS therapy after 12 months of treatment were 98% (59/60) in the D-01 study and 90% (211/235) in the D-02 study. More specifically, the percentage of subjects in a clinical trial who discontinue treatment due to adverse events is a good surrogate for the overall tolerability of the intervention being studied. In the D-01 study, 3% (2/60) of the subjects had discontinued VNS therapy for an adverse event-related reason through October 29, 2002 (by which time most continuing subjects had at least 2 years of VNS therapy). The reasons for the two discontinuations were one case of lung cancer and one death that resulted from complications following abdominal surgery. Neither of these discontinuations was related to VNS therapy or the implant procedure. In the D-02 study, 3% (8/235) of the subjects had discontinued VNS therapy for an adverse event-related reason through October 10, 2002 (by which time all continuing subjects had at least one year of VNS therapy). The reasons for these eight discontinuations included one case each of suicide, implant-related infection necessitating device removal, hoarseness, lightheadedness, post-operative pain, chest and arm pain, sudden death (of unknown cause) and worsening depression (reported by the investigator as an adverse event rather than as lack of efficacy).

These low rates of adverse event-related therapy discontinuations demonstrate that VNS therapy is very well tolerated in the population being studied.

### **3.2.4. Serious Adverse Events/Unanticipated Adverse Device Effects/Deaths**

The Sponsor did a careful review and evaluation of individual subject safety data to identify serious adverse events (regardless of potential relationship to VNS therapy), unanticipated adverse device effects, and subject deaths. It is important to note that subjects often had comorbid illnesses and almost all study subjects were also receiving antidepressant and other drugs that could have contributed to these events.

#### **3.2.4.1. Safety Considerations Specific to VNS Therapy Use in Depressed Patients**

Two specific safety concerns for the use of all antidepressant therapies are the precipitation of manic or hypomanic episodes and the effect of therapy on suicidal ideation and behavior. The findings in the VNS therapy safety database submitted in this application that relate to these two issues are summarized and discussed below.

Emergent Manic or Hypomanic Episodes: Although patients with bipolar disorder experience manic episodes as the cardinal feature of their disorder, effective antidepressant therapies themselves can precipitate a manic or hypomanic episode in these patients. Antidepressant therapies can also occasionally precipitate a manic or hypomanic episode in patients without a prior history of mania who are being treated for a major depressive episode. The emergence of hypomanic or manic reactions was evaluated in the D-01, D-02, and D-03 studies by clinical examination and by use of the Young Mania Rating Scale (YMRS). (The D-03 population is not included in the analysis below, however, because the mania data from this non-IDE study--unlike the D-03 suicide data described below--are still preliminary and unverified.) For the combined D-01 and D-02 studies, a total of 14 subjects (14/295 implanted subjects; 5%) were reported by the investigators to have at least one event of mania or hypomania or were observed to have an elevated YMRS score that emerged during the studies. Some of the subjects had multiple events. Thus, among these 14 subjects there were 13 reported events of mania or hypomania and 4 observations of a study-emergent elevation in YMRS score. Nine of the 14 subjects had a diagnosis of bipolar disorder before enrollment into the studies; the incidence of study-emergent mania or hypomania among the bipolar subjects in the combined D-01 and D-02 studies was 22% (9/41).

Since adverse events were not specifically captured in the D-04 study, the only concurrent control available to evaluate to what extent VNS therapy may have contributed to these events is the sham-control group in the D-02 study acute phase. The investigators reported two events of mania or hypomania during the acute phase. Although both events were in subjects in the VNS therapy group, one of the events occurred prior to the initiation of stimulation. To further evaluate to what extent VNS therapy may contribute to the emergence of mania or hypomania, the Sponsor did a literature review to determine the expected background incidence of such events in patients who receive antidepressant drugs. Since most of the subjects in the VNS therapy studies who experienced study-emergent mania or hypomania had bipolar disorder and since these are precisely the patients at greatest risk for treatment-emergent mania or hypomania while receiving antidepressant drugs, the literature review focused on patients with bipolar disorder. The published literature reports a wide range of incidences (but generally high) of emergent mania or hypomania in bipolar patients who are receiving antidepressant drugs. A paper by Altshuler et. al. appeared to be the most relevant publication because it examined 51 patients with treatment-refractory bipolar disorder.<sup>19</sup> This cohort is likely to be more representative of the subjects with bipolar disorders who enrolled in the D-01 and D-02 studies because the D-01 and D-02 subjects had all previously failed lithium therapy, the standard drug

used to control mania. Altshuler et. al. reported that 82% of patients switched from depression to mania during antidepressant treatment; 35% of patients had manic episodes judged to be attributable to their antidepressant drugs while for the remaining 47%, the switches were judged to be related to the natural course of their illness. Thus although comparisons with strictly historical literature controls have significant limitations, treatment-emergent mania or hypomania is widely reported to be common among patients who are treated with antidepressants, and the rate of such events among the D-01 and D-02 subjects treated with VNS therapy as an adjunct to antidepressant drugs does not appear to be higher than one would expect during treatment with the antidepressant drugs alone. To the contrary, since VNS therapy has anticonvulsant properties and anticonvulsant drugs have mood stabilizing properties, it is reasonable to speculate that VNS therapy may be an effective treatment for mania.

Emergent Suicide or Suicidal Ideation/Behavior: Although one would expect effective antidepressant therapy to reduce the risk of suicide in depressed patients, there is surprisingly little evidence in the published literature that antidepressants do so. On the other hand, published reports have raised a concern that antidepressant drugs, perhaps particularly in the pediatric population, may rarely (paradoxically) provoke suicidal ideation or behavior.<sup>20</sup> The rates of suicide and suicide attempts in the combined D-01, D-02, and D-03 safety populations were calculated on the basis of patient exposure, where a patient-year is defined as one subject receiving VNS therapy for 12 months. (The D-03 population was included in this analysis to provide the largest subject population possible, even though the D-03 study was not part of the depression studies IDE and the D-03 study data are preliminary as the study is still enrolling subjects.) These rates were then compared to the rates reported in a review by Khan et. al. of nearly 20,000 subjects participating in 45 studies comprising seven antidepressant drugs.<sup>21</sup> For the three VNS therapy studies combined, the suicide rate was 0.4% per patient year (3 events/689 patient years of exposure) and the suicide attempt rate was 3.5% per patient year (24 events/689 patient years of exposure). These rates are similar to those reported by Khan for suicide in patients receiving placebo, comparator antidepressant drugs, or investigational antidepressant drugs (0.4%, 0.7%, and 0.8% respectively) and for suicide attempt in patients receiving placebo, comparator antidepressant drugs, or investigational antidepressant drugs (2.7%, 3.4%, and 2.8% respectively).

Additionally, exploratory analyses were performed on the study D-02 and study D-04 HRSD item 3 (suicidal ideation) ratings to assess changes in suicidal ideation during VNS therapy. In one set of analyses, subjects who had an increase of two or more points in the item 3 rating were

identified. During the acute treatment phase of the D-02 study, two subjects in the VNS therapy group and three subjects in the sham-control group had such increases. Thus there was no evidence that acute VNS therapy is associated with a risk of the emergence of significant suicidal ideation compared to placebo (sham) treatment. Among subjects in the D-02 and D-04 studies who had HRSD-24 ratings after 12 months of treatment, 3% of the D-02 subjects and 2% of the D-04 subjects had increases of two or more points on item 3. These latter results, however, are difficult to interpret because the two subject groups were not well matched in the distribution of their baseline item 3 scores (D-02 subjects tended to have higher scores at baseline). Based on a slight variation of the above analysis, the percentages of subjects who had baseline scores of 0 or 1 on the suicidal ideation item of the HRSD and who showed an increase in that score (indicative of an increase in suicidal thinking/behavior) to 3 or 4 at the 12-month observation were compared for the D-02 group, the combined D-01, D-02, D-03 population, and the D-04 group. The rates of emergent suicidal thinking/behavior thus defined were similar for the three groups: 3% (4/119), 3% (4/148), and 0% (0/64) respectively.

In summary, there is no evidence from the D-01 and D-02 studies that VNS therapy poses an undue risk for precipitating hypomanic or manic episodes. A few subjects developed an increase in suicidal ideation during treatment, but these cases appear to be related to the subjects' underlying depressive disorder and not precipitated by their treatment.

#### **3.2.4.2. Other Serious Adverse Events Reported to be Possibly Related to VNS Therapy**

A serious adverse event (SAE) was defined as an event that resulted in death, was life-threatening, resulted in or prolonged hospitalization, resulted in a persistent disability, or involved a congenital anomaly. Pregnancy and cancer were also treated as SAEs. Important medical events that did not result in death, were not life threatening, or did not require hospitalization may have been considered an SAE when, based upon appropriate medical judgment, they may have jeopardized the subject and may have required medical or surgical intervention to prevent one of the outcomes listed above. SAEs were identified without regard to their relationship to treatment. In the acute phase of study D-01, there were 10 SAEs. Two SAEs were considered to be related to the implant procedure (wound infection and pain in the calf), and four SAEs were considered to be at least possibly related to stimulation (agitation, myocardial infarction in a subject with multiple risk factors for coronary artery disease, and worsened depression, [2 reports]). In the long-term phase of study D-01, there were 67 SAEs (through October 29, 2002). Of these, one SAE was considered to be related to implantation (deep vein thrombophlebitis), and eight SAEs were considered to be possibly related to stimulation (vomiting, manic reaction in a subject with

bipolar disorder, dysphoria in the same bipolar subject, worsening anxiety and depression, esophagitis, diarrhea, and suicide attempt [2 reports]). In the acute phase of study D-02, there were 30 SAEs following implantation. Of these, 11 were considered to be related to the implant procedure (wound infection, asystole, bradycardia, abnormal thinking, vocal cord paralysis, aspiration pneumonia, voice alteration, device site reaction [2 reports], acute renal failure and urinary retention). Investigators did not report any SAE to be related to stimulation. In the long-term phase of study D-02, there were 96 SAEs (through October 10, 2002). Of these, only 6 were considered to be at least possibly related to stimulation (sudden death of unknown cause, syncope [2 reports], dizziness, a manic depressive reaction in a subject with bipolar disorder, and an incident of worsening depression for which the investigator considered VNS therapy a possible but unlikely contributor).

In summary, there were relatively few reports of SAEs attributable to the VNS Therapy System implant or to VNS therapy during the D-01 and D-02 studies--further evidence of the relative safety of this therapy. The SAEs attributable to the implant procedure were all known risks of the procedure and did not occur at an unexpected rate.

#### **3.2.4.3. Serious Adverse Events of Worsening Depression**

The most commonly reported serious adverse event in the D-02 study was worsening depression. Most experts in affective disorders consider events of worsening depression as part of the underlying mood disorder and not attributable to antidepressant treatment. Therefore, not surprisingly, the investigators judged almost all the reported serious adverse events of worsening depression to be unrelated to treatment. Nonetheless, the Sponsor examined the available safety data further to determine if there was a signal that VNS therapy might paradoxically cause worsened depression. The results of these analyses support the conclusion that the occurrence of worsened depression during VNS therapy is due to the patients' underlying illness and not to VNS therapy.

Three analyses provide the support for this conclusion. First, in the acute D-02 study, there were more serious adverse events of worsening depression in the sham-control group (7) than in the treatment group (5). Second, the availability of information on hospitalizations for psychiatric illness (a reasonable proxy for serious adverse events of worsening depression) in the D-04 subjects permitted a comparison of occurrences of worsening depression between the D-02 and D-04 subjects. Calculated on a per patient-year basis, the rates of worsening depression among the D-02 and D-04 subjects were 0.293 and 0.237 respectively. Third, the percentages of subjects who had baseline scores of 0 or 1 on the depressed mood item of the HRSD and who showed an

increase in that score (indicative of an increase in depressed mood) to 3 or 4 at the 12-month observation were compared for the D-02 group, the combined D-01, D-02, D-03 population, and the D-04 group. The rates of worsening of mood thus defined were essentially the same for the three groups: 24% (5/21), 26% (6/23), and 25% (1/4) respectively. Thus the systematically collected safety data described above show that the rate of worsening depression is no greater among subjects receiving adjunctive VNS therapy than it is among subjects receiving placebo (sham VNS), and it is similar to the rate among subjects receiving antidepressant drugs without VNS therapy.

#### **3.2.4.4. Unanticipated Adverse Device Effects (UADEs)**

No events in the D-01 study met criteria for an UADE. Two events in the D-02 study met criteria for an UADE. Both these events were non-specific complications of surgery related to the implant procedure and occurred prior to the initiation of stimulation. The first of these UADEs was an episode of acute renal failure thought to be secondary to antibiotic administration. The second UADE was an episode of altered mental status thought to be due to perioperative narcotic administration.

#### **3.2.4.5. Deaths**

One death occurred during the D-01 study. The subject died of sepsis related to abdominal surgery. A second D-01 study subject died of lung cancer after withdrawing from the study; the subject's VNS Therapy System had already been explanted. Four deaths occurred during the D-02 study; one of these occurred prior to implantation. The remaining three deaths were a suicide, a sudden death of unknown cause, and a subject with multi-organ failure. No direct role for VNS therapy was evident in any of these deaths.

#### **3.2.5. Physical and Neurological Examinations, Vital Signs Assessments, Holter Monitor Recordings**

Investigators performed physical and neurological examinations in D-01 and D-02 study subjects at baseline and periodically during treatment. The post-baseline examinations were essentially unremarkable except for the presence of healed surgical implant scars and the occasional presence of voice alteration. For the D-02 study, changes from baseline in blood pressure, heart rate, respiratory rate, and weight were plotted graphically. The changes generally conformed to a symmetrical distribution, suggesting an absence of effect associated with VNS therapy. Holter monitor recordings were obtained as part of the D-01 study. The recordings did not show any evidence of significant changes after VNS therapy was initiated.

### 3.3. Device Performance

The VNS Therapy System performed as expected in the D-01 and D-02 studies. The most significant device performance issue detected in the two studies was an incidence of a high impedance reading in the D-02 study that was due to a lead break. The lead was replaced. (A second incidence of a high impedance reading was determined on further investigation to be an error and did not require further action.)

### 3.4. Safety Conclusions

The results from the depression studies reconfirm the valid scientific evidence for safety that was established in the Sponsor's Premarket Approval application for the VNS therapy epilepsy indication (P#970003).

- ?? The safety profile of VNS therapy observed in the studies submitted in this application is consistent with the known safety profile of VNS therapy established by the epilepsy IDE trials and over 56,000 patient-years of commercial use.
- ?? VNS therapy was associated with few safety concerns and was well tolerated by subjects in the D-01 and D-02 clinical trials.
- ?? There was no indication from the data collected in the D-01 and D-02 studies to suggest that VNS therapy causes treatment-emergent mania or hypomania or an increase in suicidal ideation or behavior.
- ?? The most commonly reported events judged to be related to the implant procedure were device site pain, device site reaction, incision pain, dysphagia, hypesthesia, pharyngitis, voice alteration, and incision site reaction. These events generally resolved within 30 days, although hypesthesia and voice alteration were more likely to persist than were the other events.
- ?? The most commonly reported acute events judged to be related to stimulation were neck pain, dyspepsia, dysphagia, vomiting, paresthesia, increased cough, dyspnea, laryngismus, and voice alteration. Reports of stimulation-related events tended to diminish markedly within 60 days except for laryngismus and voice alteration, which tended to be more persistent.
- ?? VNS therapy is generally very well tolerated as confirmed by a low rate of study discontinuations due to adverse events (3% through the D-01 and D-02 data cutoff dates, which allowed all continuing subjects to receive at least 12 months of VNS therapy).
- ?? Few SAEs attributable either to VNS Therapy System implantation or VNS therapy were observed in these studies. During these studies, four deaths occurred in subjects receiving VNS. (One additional death occurred prior to implant and one occurred after explant.) The four deaths included one suicide and one sudden death of unknown case, but no deaths directly attributable to VNS therapy.
- ?? Physical and neurological examinations, vital signs assessments, and Holter monitor recordings during VNS therapy did not demonstrate any significant adverse effects from VNS therapy.
- ?? The VNS Therapy System components performed according to expectations.

## **4. Risks, Benefits, Clinical Utility, and Overall Conclusions**

### **4.1. Risks Associated with the Use of VNS Therapy in the Treatment of Depression**

The potential risks associated with the use of VNS therapy in the treatment of epilepsy that is refractory to antiepileptic medications have been well-established during premarket approval clinical trials that included 454 subjects implanted with the VNS Therapy System and during commercial experience that includes VNS Therapy System implants in more than 22,000 patients comprising over 56,000 patient-years of experience. The clinical safety data submitted in this application demonstrate that the risk profile established during the epilepsy trials and commercial experience is not altered in any substantial way when VNS therapy is used to treat patients with treatment-resistant depression.

The potential risks associated with the use of VNS therapy are well documented in the current Physician's Manual for the VNS Therapy Pulse Model 102/102R Generator. They include the following:

- ?? the potential for human tissue damage if diathermy or full body magnetic resonance imaging is performed in a patient implanted with the VNS Therapy System
- ?? the potential for degenerative nerve damage to the vagus nerve if it is stimulated excessively (ie, if stimulation on time exceeds off time)
- ?? aspiration related to stimulation-induced impairment of swallowing, especially in the presence of predisposing factors
- ?? painful stimulation if the device malfunctions
- ?? increased apneic events in patients with obstructive sleep apnea
- ?? dyspnea, especially in the presence of obstructive airway disease
- ?? transient bradycardia or asystole, especially during initial intraoperative lead testing
- ?? surgical risks associated with the implant procedure, principally infection, vocal cord dysfunction due to manipulation of the vagus nerve, and other nerve damage
- ?? a variety of non-serious side effects associated with stimulation, most commonly neck pain, dyspepsia, dysphagia, vomiting, paresthesia, increased cough, dyspnea, laryngismus, and voice alteration.

### **4.2. Benefits/Clinical Utility of VNS Therapy in the Treatment of Depression**

The results presented in this application are derived from clinical trials in which VNS therapy was used as an adjunctive long-term treatment for chronic or recurrent depression in subjects

experiencing a major depressive episode that had not had an adequate response to two or more adequate antidepressant treatments.

Under those conditions of use, the clinical trial results presented in this application support the following benefit/clinical utility profile. VNS therapy can be expected to produce:

- ?? a response (ie, at least a 50% improvement in depressive symptoms) in at least 15% to 17% of patients after 10 weeks of treatment
- ?? a response in at least 22% to 37% of patients after 12 months of treatment
- ?? a complete response (usually referred to as a remission) in at least 15 to 17% of patients after 12 months of treatment
- ?? a sustained response in 13% to 27% of patients during 12 months of treatment
- ?? successful maintenance of the initial improvement in a high percentage of patients (eg, 73% to 77% of patients who had meaningful or greater benefit after 3 months of treatment maintained at least meaningful benefit after 12 months of treatment)

Additionally, the unique device-based non-systemic mechanism of action associated with VNS therapy provides these benefits:

- ?? a very well-tolerated treatment as indicated by the high continuation rates for VNS therapy in the D-01 and D-02 studies after 12 months of treatment (90% to 98%) and the low rate of adverse event-related study discontinuations through 12 months or more in these studies (3%)
- ?? adverse effects essentially limited to those related to stimulation of the vagus nerve, ie, an absence of the systemic effects associated with drug therapy; many of the common adverse effects only occur when stimulation of the vagus nerve is actually occurring
- ?? the ability to halt acute stimulation-related adverse effects immediately through the use of magnet deactivation of stimulation
- ?? an absence of adverse cognitive and psychomotor effects observed with antidepressant drugs and ECT
- ?? absence of the overdose toxicity observed with antidepressant drugs
- ?? favorable findings in animal reproductive studies
- ?? ability to add VNS therapy to antidepressant drug therapy without producing drug-drug interactions
- ?? high treatment compliance because VNS therapy is programmed to work automatically without the need for patient action (ie, no pills to take)

#### **4.2.1. Introduction to Patient Video**

This application includes a brief videotape of four subjects from the D-01 and D-02 studies to provide a patient perspective on the benefits and clinical utility of VNS therapy in the treatment of depression.

#### **4.3. Overall Conclusions**

The results of the clinical studies conducted under IDE G980099 (studies D-01 and D-02) provide valid scientific evidence to support the safety and effectiveness of the adjunctive use of VNS therapy to treat patients with chronic or recurrent depression over the age of 18 who are experiencing a major depressive episode that is resistant to two or more adequate antidepressant therapies. Valid scientific evidence of effectiveness is derived from a comparison of depression symptom ratings in subjects who received VNS therapy plus standard-of-care treatment over 12 months with ratings in subjects who received only standard-of-care treatment over 12 months. Subjects in the VNS therapy group showed significantly more improvement in depressive symptoms than did subjects in the standard-of-care treatment group ( $p < 0.001$  in the primary analysis). Twenty two percent (22%) to 37% of the VNS therapy group (depending on the rating scale used) achieved a response during 12 months of treatment. The effectiveness of VNS therapy in the treatment of depression is further supported by the finding that VNS therapy produced higher numerical response rates than did sham stimulation over 10 weeks of acute treatment (although the difference did not reach statistical significance in the primary analysis) and by the finding that acute response to VNS therapy was maintained in over 60% of subjects during 12 months of treatment. Results from safety assessments in these two studies show that VNS therapy was well-tolerated and was associated with few significant safety concerns. Overall, the safety profile of VNS therapy established in epilepsy clinical trials and commercial use does not appear to be altered when VNS therapy is used in patients with treatment-resistant depression. In summary, the VNS Therapy System is both safe and effective for the proposed indication when used in accordance with the directions for use found in the VNS Therapy System Physician's Manual.

## References

1. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003; 289:3095-3105.
2. Murray CJ, Lopez AD. Evidence-based health policy lessons from the Global Burden of Disease Study. *Science*. 1996; 274:740-743.
3. Thase ME, Rush AJ in Bloom SE, Kupfer DJ (eds): *Psychopharmacology: The Fourth Generation of Progress*. New York; Raven Press Ltd. 1995; 1081-1097.
4. Fawcett J. Antidepressants: Partial response in chronic depression. *Br J Psychiatry*. 1994; 165:37-41.
5. Paykel ES, Ramana R, Cooper Z, et al. Residual symptoms after partial remission: an important outcome in depression. *Psychological Medicine*. 1995; 25:171-1180.
6. Miller IW, Keitner GI, Schatzberg AF, et al. Psychosocial functioning before and after treatment with sertraline or imipramine. *J Clin Psychiatry*. 1998; 59:608-619.
7. Harden CL, Pulver MC, Ravdin LD, et al. A pilot study of mood in epilepsy patients treated with vagus nerve stimulation. *Epilepsy Behav*. 2001; 1:93-99.
8. Elger G, Hoppe C, Falkai P, et al. Vagus nerve stimulation is associated with mood improvements in epilepsy patients. *Epilepsy Res*. 2000; 42:203-210.
9. Henry TR. Therapeutic mechanisms of vagus nerve stimulation. *Neurology*. 2002; 59(Suppl 4): S3-S14.
10. Rush AJ, Trivedi M, Carmody TJ, et al. One-Year clinical outcomes of depressed public sector outpatients: A benchmark for subsequent studies. In press, *Biological Psychiatry*.
11. Thase ME, Rush AJ. Treatment-Resistant Depression. *Psychopharmacology: The Fourth Generation of Progress*, edited by Floyd E. Bloom and David J. Kupfer. 1995; 1081-1097.
12. Quitkin FM, Rabkin JD, Markowitz JM, et al. Use of pattern analysis to identify true drug response. A Replication. *Arch Gen Psychiatry*. 1987; 44:259-264.
13. Keller MB, Boland RJ. Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression. *Biol Psychiatry*. 1998; 44:348-360.
14. Nierenberg AA, Feighner JP, Rudolph R, et al. Venlafaxine for Treatment-Resistant Unipolar Depression. *J Clin Psychopharmacol* 1994; 14:419-423.
15. Sackeim HA, Prudic J, Devanand DP, et al. A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Arch Gen Psychiatry*. 2000; 57:425-434.
16. Sackeim HA, Haskett RF, Mulsant BH, et al. Continuation Pharmacotherapy in the prevention of relapse following electroconvulsive therapy: a randomized controlled trial. *JAMA*. 2001; 285:1299-1307.
17. Prudic J, Olfson M, Marcus SC, et al. The effectiveness of electroconvulsive therapy in community settings. *JAMA*. 2003; in press.
18. U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research. Depression in Primary Care: Volume 2. Treatment of Major Depression. *Clinical Practice Guideline Number 5*. 47-51.

19. Altshuler LL, Post RM, Leverich GS, et al. Antidepressant-Induced mania and cycle acceleration: a controversy revisited. *Am J Psychiatry*. 1995; 152:1130-1138.
20. Varley CK. Psychopharmacological treatment of major depressive disorder in children and adolescents. *JAMA*. 2003; 290(8): 1091-1093.
21. Khan A, Warner HA, Brown WA. Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials. *Arch Gen Psychiatry*. 2000; 57:311-317.